VOLUME 25 NOVEMBER, 1937 NUMBER 11

PROCEEDINGS
of
The Institute of Radio Engineers

Application Blank for Associate Membership on Page IX

Institute of Radio Engineers
Forthcoming Meetings

ROCHESTER FALL MEETING
Sagamore Hotel
Rochester, New York
November 8, 9, and 10, 1937

DETROIT SECTION
November 19, 1937

LOS ANGELES SECTION
November 16, 1937

MONTREAL SECTION
November 10, 1937

NEW YORK MEETING
November 3, 1937
December 1, 1937

PHILADELPHIA SECTION
November 4, 1937
December 2, 1937

TORONTO SECTION
November 8, 1937

WASHINGTON SECTION
November 8, 1937

October Meeting of the Board of Directors

The regular meeting of the Board of Directors was held on October 6 in the Institute office and attended by H. H. Beverage, president; Stuart Ballantine, Ralph Bown, Alfred N. Goldsmith, Virgil M. Graham, L. C. F. Horle, C. M. Jansky, Jr., C. B. Jolliffe, Haraden Pratt, H. M. Turner, and H. P. Westman, secretary.

Thirty-three new Associates, two Juniors, and three Students were elected to membership.

It was agreed that a two-day program be prepared for the joint meeting of the Institute and the American Section of the International Scientific Radio Union in Washington, D.C.

Approval was granted of a letter to be forwarded to President Roosevelt conveying our satisfaction in the appointment of T. A. M. Craven as a member of the Federal Communications Commission. Commander Craven is a Fellow of the Institute and a member of the Board of Directors.

An invitation to appoint a representative of the Institute on a committee to revise the radio portion of the National Electric Safety Code was accepted.

The Standards Report prepared by the Technical Committee on Electronics and approved by the Standards Committee was adopted. Similarly, a report on the testing of loud-speakers prepared by the Technical Committee on Electroacoustics was adopted. These reports are now being prepared for printing and copies will be distributed to all members of the Institute upon completion.

An invitation to continue our sponsorship and representation in the preparation for the Second National Conference on Educational Broadcasting which will be held in Chicago, Illinois, on November 29 and 30, and December 1, 1937 was accepted. C. M. Jansky, Jr., was designated as our representative.

Rochester Fall Meeting

The Rochester Fall Meeting for 1937 will be held in the Sagamore Hotel in Rochester, N.Y., on November 8, 9, and 10. The program is as follows:

MONDAY, NOVEMBER 8

9:00 A.M. Registration
Inspection of Exhibits

10:00 A.M. **Technical Session**
"Parallel Resonance Methods for Measurement of High Impedances at High Frequencies," by D. B. Sinclair, General Radio Company.
"Report of RMA Television Transmission Frequencies and Standards," by A. F. Murray, Philco Radio and Television Corporation.
"Vibrational Tube Analysis" (with demonstration), by A. B. Oxley, RCA Victor Company (Canada).

12:30 P.M. Group Luncheon

2:00 P.M. **Technical Session**
"The Problem of Synchronization in Cathode-Ray Television," by F. J. Bingley, Philco Radio and Television Corporation.
"New High Efficiency Modulation System," by R. B. Dome, General Electric Company.

4:00 P.M. Inspection of Exhibits
RMA Committee Meetings

6:30 P.M. Group Dinner

7:45 P.M. **Technical Session**
"Specification of Screen Color of Cathode-Ray Tubes," by G. A. Fink and R. M. Bowie, Hygrade Sylvania Corporation.
"Figure of Merit for Television Performance," by A. V. Bedford, RCA Manufacturing Company, RCA Victor Division.

TUESDAY, NOVEMBER 9

9:00 A.M. Registration
Exhibits Open

9:30 A.M. **Technical Session**
"Direct Viewing Type Cathode-Ray Tube for Large Television Images," by I. G. Maloff, RCA Manufacturing Company, RCA Victor Division.
"Stabilization of Oscillators," by C. E. Granqvist, Stockholm, Sweden.
"A Unique Method of Modulation for High Fidelity Television Transmitters" (with demonstration), by William N. Parker, Philco Radio and Television Corporation.

12:30 P.M. Group Luncheon

2:00 P.M. **Technical Session**
"New Projects of the RMA Engineering Division," by L. C. F. Horle, RMA Engineering Division.
"Measurement of Characteristics of Automobile Antennas," by H. Lyman, Philco Radio and Television Corporation. Discussion by H. C. Forbes, Colonial Radio Corporation, and D. E. Foster, RCA License Laboratory.

2:00 P.M. **Physicists' Session**
"Space-Charge Limitation on the Focus of Electron Beams," by B. J. Thompson and L. B. Headrick, RCA Manufacturing Company, RCA Radiotron Division.
"Negative Ion Components of the Cathode Ray," by C. H. Bachman and C. W. Carnahan, Hygrade Sylvania Corporation.

4:00 P.M. Inspection of Exhibits
RMA Committee Meetings

6:30 P.M. Stag Banquet

WEDNESDAY, NOVEMBER 10

9:00 A.M. Exhibits Open

9:30 A.M. **Technical Session**

"The Monoscope," by C. E. Burnett, RCA Manufacturing Company, RCA Radiotron Division.

"Further Data on Inverse Feed-Back Amplifiers," by C. B. Fisher, Northern Electric Company.

"Teledynamic Control by Selective Ionization and the Application to Radio Receivers" (with demonstration) by S. W. Seeley, H. B. Deal, and C. W. Kimball, RCA License Laboratory.

12:30 P.M. Group Luncheon

2:00 P.M. **Technical Session**

"Stability of Wide-Band Amplifiers," by E. H. B. Bartelink, General Electric Company.

"An Audio Curve Tracer," (with demonstration) by J. B. Sherman, RCA Manufacturing Company, RCA Radiotron Division.

4:00 P.M. Exhibits Close
RMA Committee Meetings

Broadcast Engineering Conference

A Broadcast Engineering Conference will be conducted by the Ohio State University during the period from February 7 to 19. In addition to work given by the staff, prominent speakers will discuss such subjects as broadcast antenna design, high powered amplifiers, modulation and distortion measurements, studio acoustics, and ultra-high-frequency propagation. The full program will appear in a forthcoming issue of the PROCEEDINGS. Those wishing more information may obtain it from the Director, W. L. Everitt, Department of Electrical Engineering, Ohio State University.

Committee Work

CONSTITUTION AND LAWS COMMITTEE

H. M. Turner, chairman, Austin Bailey, J. D. Crawford, assistant secretary, and H. P. Westman, secretary, attended a meeting of the Constitution and Laws Committee which was held in the Institute office on September 24. The committee continued its work on the revision of the Institute constitution.

Technical Committee on Electronics

Subcommittee on Small High Vacuum Tubes

The Subcommittee on Small High Vacuum Tubes operating under the Technical Committee on Electronics met in the Institute office on September 30. Those present were P. T. Weeks, chairman; L. E. Barton, R. S. Burnap, C. B. Upp, J. D. Crawford, assistant secretary, and H. P. Westman, secretary. The committee devoted its time to an analysis of its program of operation looking toward the development of new material for inclusion in a future issue of a standards report to be issued by the Technical Committee on Electronics.

Institute Meetings

Atlanta Section

On September 9 the Atlanta Section met at the Atlanta Athletic Club with P. C. Bangs, vice chairman, presiding and thirty in attendance. W. G. Cady, professor of physics of Wesleyan University presented a paper on "The How and Why of Obliquely Cut Piezo Crystals." Dr. Cady first presented a historical outline of piezo electricity and the early developments of the use of piezoelectric crystals for radio purposes. The importance of the angles at which plates are cut from the original crystals was then shown and the effect on the temperature coefficient in relation to frequency stressed. He closed his paper with a description of some work on Rochelle salt crystals and described a method of modulating light rays by passing them through a liquid in which vibrations of the desired period were set up by means of oscillating crystals. He included in his paper observations made during his recent visits to laboratories in France and England.

Cincinnati Section

George Platts, chairman, presided at the September 21 meeting of the Cincinnati Section which was attended by forty-five and held in the University of Cincinnati.

A paper on "Modern Receiver Design" was presented by D. E. Foster of the RCA License Laboratory. In it he discussed the procedure involved in designing a broadcast receiver from the embryonic idea of the sales department to the final sample submitted to the production department. An outline was given of the engineering problems, cost considerations, factors governing design, layout difficulties, and other points of importance in the creation of a modern receiver. The paper was discussed by Messrs. Barbulesco, Kilgour, and Osterbrock.

Cleveland Section

A meeting of the Cleveland Section was held on May 28 at WHK Studios with R. A. Fox, chairman, presiding. There were nineteen present.

C. B. Aiken, associate professor of electrical engineering at Purdue University presented a paper on "High-Q Tuned Coupled Circuits." This paper was based on the one published in the Proceedings for February, 1937.

Connecticut Valley Section

F. H. Scheer, chairman, presided at the May 20 meeting of the Connecticut Valley Section at which thirty-five were present. The meeting was held in Springfield, Mass.

A paper on "Vacuum Tubes for use at Ultra-High Frequencies" was presented by A. L. Samuel of the technical staff of the Bell Telephone Laboratories. The subject was introduced by pointing out the overcrowding in the radio spectrum which has made necessary its expansion in the high-frequency region by the development of equipment of commercial usefulness. He described a double pentode amplifying tube which employs two sets of elements mounted within the same envelope, is mechanically constructed to permit accurate maintenance of extremely small spacings between elements, employs very short leads which are not common to two or more elements, provides shielding between the input and output terminals, and establishes a ground within the envelope for the screen and suppressor grids. This type of tube overcomes the difficulties encountered by conventional construction in which interaction between input and output circuits is caused by excessive grid-plate capacitance, impedance in the form of lead inductance, and long electron transit time. Graphs were given of the characteristics of the new tube and methods of using it in amplifying circuits at frequencies from 150 to 300 megacycles were described.

Detroit Section

The Detroit Section held a meeting on September 17 in the Conference Room of the Detroit *News*. R. L. Davis, chairman, presided and there were thirty-two present.

"Recent Magnetic Theory" was the subject of a paper by G. P. Brewington, professor of physics of Lawrence Institute of Technology. In it Dr. Brewington pointed out that the magnetic properties displayed by iron and certain other materials are related to their atomic structure. The paper was devoted largely to a description of some of the simpler relationships between the positions of the atoms in the crystal

and the magnetic properties of the material. The principle magnetic materials discussed were iron, chromium, and nickel.

Emporium Section

M. I. Kahl, chairman, presided at the September 16 meeting of the Emporium Section which was held in the American Legion Rooms and attended by thirty-seven.

A paper on "Design and Manufacture of Adjustable Resistance Units" was presented by H. A. Williams, superintendent of the Stackpole Carbon Company volume control plant. He described first the mechanical and electrical properties needed in adjustable resistance units such as volume controls. The size of the unit is of necessity a compromise between the two requirements. The wide range of materials used in these units was described. Although subject to relatively small movement, lubrication is an important factor in the life of these units. Because of the varying characteristics demanded by set manufacturers, the units must be manufactured on equipment which is not completely automatic.

The molded type is considered to be inherently noisy and difficult to manufacture to give other than linear characteristics. The disk type although entirely satisfactory is expensive to manufacture and employs between fifty and sixty parts. It consists of alternate disks of carbon and metal, the wiping fingers of the control riding on the metal. A stepped curve results from this type of construction. The film type is quiet in operation and may be produced to give a wide range of resistance variation with rotation of the control. In this type the conductor is a film of paint generally about two hundredths of an inch in thickness and sprayed on an insulated base. Details of the manufacture of these units were given. The paper was concluded with some general applications and the demands they make on the electrical and mechanical characteristics of the units. An outline was given of standards of mechanical strength, allowable humidity variation, stray capacitance, and noise generated in the operation of the devices. The paper was discussed by Messrs. Bowie, Campbell, Dehlinger, Kahl, and West.

New Orleans

The New Orleans Section accepted an invitation to attend a meeting sponsored by the New Orleans Amateur Radio Club on August 24, in the Association of Commerce building at which a paper was given by C. L. Reinartz, research engineer of the RCA Manufacturing Company, which covered the design of tank circuits in radio transmitters, half-wave "curved" antennas, and vacuum tube voltmeters. G. H.

Peirce, secretary of the Institute section, presided and there were thirty-eight present. There were about twelve present at the dinner which preceded the meeting.

San Francisco Section

Two meetings of the San Francisco Section were held during September in the auditorium of the Pacific Telephone and Telegraph Company. The first, on the eighth, was attended by fifty-four and presided over by V. J. Freiermuth, chairman. President Beverage was president at this meeting and spoke a few words of greeting to those in attendance.

W. G. Wagener, tube design engineer of the RCA Manufacturing Company, presented a paper on "Ultra-High-Frequency Tubes." The author described tubes used as ultra-high-frequency oscillators and amplifiers and showed curves giving the maximum power output as related to the frequency of operation for many well known tubes. The limitation in tube output caused by electron transit time was discussed in detail. A description was then given of two new tubes designed specifically for ultra-high-frequency operation. These are the RCA 887 and 888 and are capable of substantial outputs at wavelengths of the order of 125 centimeters.

The meeting on the twenty-second was attended by twenty-five and presided over by Noel Eldred, vice chairman. A discussion of the paper "The Shunt-Excited Antenna" by J. F. Morrison and P. H. Smith which appeared in the June, 1937, Proceedings was led by P. A. Ekstrand of the engineering staff of Heintz and Kaufman.

A paper on "Characteristics of American Broadcast Receivers as Related to the Power and Frequency of Transmitters" was presented by James Sharp of the Pacific Telephone and Telegraph Company.

Washington Section

On September 13 the Washington Section met in the Potomac Electric Power Company auditorium. There were seventy-five present and W. B. Burgess, chairman, presided.

"Recent Developments in Diversity Receiving Equipment" was the subject of a paper by J. B. Moore, an engineer for R.C.A. Communications. He described in considerable detail the theory of the diversity system and its advantages in providing reliable communication under adverse conditions of fading. A description was given of the operating characteristics and circuit features of the system and the problems encountered in the design of equipment.

Correction

K. A. Norton has brought to the attention of the editors the following errors which appeared in his papers, "The Physical Reality of Space and Surface Waves in the Radiation Field of Radio Antennas" and "The Propagation of Radio Waves over the Surface of the Earth and in the Upper Atmosphere, Part II," published in the September, 1937, issue of the PROCEEDINGS:

Page 1195. The upper limit on the integral in equation (8) should be ∞.

Page 1200. Equation (15) should read

$$R_h = \frac{\sqrt{1 - u^2 \cos^2 \psi} - u \sin \psi}{\sqrt{1 - u^2 \cos^2 \psi} + u \sin \psi}$$

Page 1207. The signs of the first six terms (appearing in the first two rows) of equation (15) should be negative.

Pages 1227 and 1228. The factor $[-\cos((\pi/2)\cos\phi)/\sin^2\phi]$ was omitted from equations (101) through (106).

Page 1229. Equation (110) . . . E_z^v should read $|E_z^v|$.

Proceedings of the Institute of Radio Engineers
Volume 25, Number 11 *November, 1937*

TECHNICAL PAPERS

AN ELECTRODYNAMIC AMMETER FOR USE AT FREQUENCIES FROM ONE TO ONE HUNDRED MEGACYLES*

By
H. M. Turner and P. C. Michel
(Yale University, New Haven, Connecticut)

Summary—*Although thermoammeters are used universally for measuring current at high frequencies, there has been until recently no independent method of checking their accuracy. The electrodynamic ammeter was developed to meet this need by providing a standard, based on an entirely different principle of operation. It gives an absolute method for the measurement of current at frequencies above a megacycle. A description and theory of the instrument is presented.*

Description of Instrument

THIS ammeter is designed for measuring current of the order of one to five amperes at frequencies from one to one hundred megacycles and higher. It may be thought of as a special type air-cored transformer consisting of a single turn primary mounted in a vertical plane, which carries the current to be measured, and a smaller closed turn secondary, vertically suspended by a quartz fiber. The secondary when angularly displaced from the position of zero coupling is acted upon by forces which produce torsional oscillations about its vertical diameter. The frequency of these mechanical oscillations is directly proportional to the magnitude of the primary current but independent of its electrical frequency. A photograph of the instrument, in series with a thermoammeter, arranged for measuring the current at the end of a concentric transmission line extending to the left, is shown in Fig. 1.

The secondary turn, or ring, and the quartz suspension are shielded from external air currents by a glass tube terminating in a spherical bulb, a portion of which is cut away to permit the insertion of the ring after which it is replaced and sealed but not exhausted. To prevent excessive internal viscous damping there should be a clearance of at least half a centimeter between the ring and the bulb. In a particular instru-

* Decimal classification: R242.1. Original manuscript received by the Institute, August 19, 1936; revised manuscript received by the Institute, January 4, 1937; August 2, 1937. Presented before a joint meeting of the Institute of Radio Engineers and the American Section of the International Scientific Radio Union, Washington, April 1934, and at the General Assembly of the International Scientific Radio Union, London, September 1934.

ment having a ring of No. 26 wire, two centimeters in diameter, in a bulb of 2.4 centimeters inside diameter, the damping was sufficiently great to prevent oscillations; however, when a bulb of 3.2 centimeters was substituted the operation of the instrument was normal. Viscous damping does not affect the accuracy other than through reducing the time over which oscillations may be counted. With forty to sixty seconds available, the results are entirely satisfactory. When small currents are to be measured it would be advantageous to establish a fairly high vacuum within the bulb and thus reduce the air damping. This would involve a different tube design.

Fig. 1—Electrodynamic ammeter and thermoammeter in series for measuring current at end of concentric transmission line.

Practically it would be simpler to enclose both primary and secondary in a shield but in an earlier model where this was done a peculiar phenomenon was observed. In the normal use of the instrument the torsional oscillations die out in a minute or so with the secondary coming to rest in a position perpendicular to the primary where the coupling is zero. Now, opening the primary should have no effect whatever on the secondary, which was true in this case at the moment of opening, but after an interval of about a minute the ring would swing through quite an angle and then gradually drift back to its original position. An investigation showed it to be caused by the perturbance of the air within the shield, during cooling, acting on the silk suspension which was then used. This was rather surprising in view of the fact that there was less than one-tenth watt dissipated in the primary and the radiating surface was quite large. However, by decreasing the resistance of the primary coil sufficiently this effect was eliminated. It did not affect

the accuracy but nevertheless was annoying. By placing the primary outside the shield, as previously discussed, the heat from it was transferred to the surrounding space instead of accumulating within the shield. This eliminated the difficulty.

It was found that quartz fiber suspensions were more satisfactory than silk ones and so they were later used exclusively. The quartz suspensions varied from two to twenty microns in diameter and from twelve to twenty centimeters in length. The rings were usually of fourteen-gauge copper and of one to four centimeters in diameter. The natural period of oscillation of the secondary member, that is, with no current in the primary, is of the order of forty seconds, showing the relatively small restoring force of the suspension. Due to this fact it is not necessary to make the free position of the secondary exactly perpendicular to the plane of the primary, but only approximately so. In fact, the normal procedure in using the instrument is to orient the suspension head so that the free position of the ring is displaced ten or twelve degrees from a vertical plane perpendicular to the plane of the primary and passing through its center so oscillations of proper amplitude are set up as soon as current flows in the primary. By keeping the swing within twelve degrees either side of the perpendicular, which is adequate for measuring purposes, the electrical reaction of the secondary on the primary is negligible.

Fig. 2—Resultant field produced by currents 180 degrees out of time phase.

Theory of Operation

The secondary induced voltage lags ninety degrees behind the primary current, and, at frequencies where the instrument is used, the resulting secondary current lags ninety degrees behind the voltage, thus the two currents are one hundred and eighty degrees out of time phase. With the secondary, or smaller coil, in any position with respect to the primary, such as shown in Fig. 2, consider a plan view of the superposed fields in a horizontal plane through the coil centers, which

illustrates the action of the high-frequency ammeter at any instant of time. As is evident from the crowding of the lines of force at the extreme right and the extreme left of the secondary, where it passes through the paper, the freely suspended secondary will move in a counterclockwise direction and interlink fewer lines of force from the primary, that is, towards a position perpendicular to the plane of the primary. When the secondary reaches a position of zero coupling, motion would cease were it not for the kinetic energy of the moving coil carrying it beyond into a region of opposing torque which increases in magnitude as the motion continues until it finally brings the coil to rest. This torque, due to the interaction of the field, then reverses the direction of motion and this cycle of events is repeated thus producing mechanical oscillations.

Before developing a mathematical expression for the current to be measured certain fundamental concepts will be outlined. Consider, for a moment, a circuit in which a constant current I flows through two air-cored coils in series, inductively aiding. The total energy stored in the magnetic form is $1/2(L_1+L_2+2M)I^2$ of which MI^2 is in the mutual field. Let one coil be given a linear displacement, dx, increasing the mutual inductance by dM in time dt. Since the resistance of the circuit will have no effect on the work done in moving the coil, it will be omitted from consideration in the interest of simplicity. To keep the current constant during the motion requires that the impressed electromotive force be increased by an amount equal to that induced by the change in M. A common expression for induced electromotive force is $e=L(di/dt)$ where L is constant and i variable. A more general form is $e=d/dt(Li)$. In the case under consideration the current and self-inductances of the two coils are constant and the mutual inductance is the variable, so the expression becomes $e=I(dL/dt)=2I(dM/dt)$, since $L=L_1+L_2+2M$. The work done in moving the coil may now be calculated.

Work $=Ie\,dt=2\,I^2dM$, the result being expressed in ergs or joules depending upon whether electromagnetic or practical units are used. It will be observed that the work is independent of the time taken to move the coil. Of the $2I^2dM$ units of work, I^2dM represents the increase in the energy stored in the mutual field, which follows directly from the last paragraph, and the remaining I^2dM is the mechanical work done in moving the coil, that is, $f\,dx=I^2dM$ or $f=I^2dM/dx$. For the case of angular displacement, as in the ammeter, replacing dx by its equivalent $R\,d\theta$ and expressing results in terms of torque gives

$$T = I^2\frac{dM}{d\theta}\text{ dyne-centimeters.}$$

The energy in the mutual field, MI^2, is obtained by multiplying M by the product of the current in the two coils, or more generally Mi_1i_2, regardless of the manner of connection. The instantaneous torque is then given by the expression

$$T = i_1 i_2 \frac{dM}{d\theta} \cdot$$

Having outlined the fundamental concepts, an expression for the current to be measured will now be developed in terms of the instrument constants and the mechanical frequency of torsional oscillation. The following symbols will be used, all units being in the electromagnetic system.

I = the effective value of current being measured,

$i = \sqrt{2}\, I \sin \omega t$, the instantaneous value,

f = frequency of the current,

F = mechanical frequency of torsional oscillations of the suspended ring in the field of the current being measured,

M = mutual inductance between primary and secondary,

e_s = voltage induced in the secondary by the current i

$= -\dfrac{d}{dt}(Mi) = -M\dfrac{di}{dt} - i\dfrac{dM}{dt} = -M\dfrac{di}{dt}$ to one part in a million for frequencies above 10^6 where the ring may be considered stationary during one cycle of the current as a result of which M is constant over the cycles,

$= -\sqrt{2}\,\omega M I \cos \omega t.$

i_s = current in the secondary ring,

$= -\sqrt{2}\,\dfrac{MI}{L_s} \sin \omega t$, to one part in a thousand for frequencies above 10^6 where the ring power factor is practically zero,

L_s = inductance of the ring,

J_s = moment of inertia of the ring about its vertical diameter,

θ = angular displacement of the ring, from position of zero coupling, in radians,

T' = instantaneous torque in dyne-centimeters,

T = average torque in dyne-centimeters.

The instantaneous torque on the ring, produced by the current in the two coils of the ammeter is in accordance with the relation previously developed

$$T' = i i_s \frac{dM}{d\theta} = -\frac{2I^2}{L_s} \sin^2 \omega t\, M \frac{dM}{d\theta} \cdot$$

The average torque is found by integrating over a current cycle and dividing by the corresponding interval. Since the frequency of the current is so great compared with the mechanical frequency of oscillation of the ring, $dM/d\theta$ is constant over the cycle, and the average torque is

$$T = -\frac{I^2}{L_s} M \frac{dM}{d\theta} \cdot$$

Strictly speaking, the value of L_s varies slightly with the frequency, due to skin effect, but approaches closely to its limiting value for frequencies above one megacycle. If one uses, without correction, the limiting value of L_s obtained at infinite frequency the error will be $1\frac{1}{4}$ per cent at one megacycle and less than $\frac{1}{4}$ per cent above twenty megacycles, regardless of the wave form of the current.

For small values of θ the magnitude of $MdM/d\theta$ is proportional to θ. If the quantity m be defined by the equation

$$M\, dM/d\theta = m^2\theta,$$

then m is a calculable function of the primary and secondary turn diameters and their horizontal separation. For one configuration used m is constant to within $\frac{1}{4}$ per cent for values of θ up to one-fifth radian, or about twelve degrees, either side of the position of zero coupling.

The restoring torque, which becomes $-I^2m^2\theta/L_s$, is opposed by the inertial reaction of the ring, $J_s d^2\theta/dt^2$; by Newton's third law it is equal thereto, neglecting viscous air damping and suspension torque, both of which are experimentally as well as theoretically proved negligible, or less than $\frac{1}{4}$ per cent in the operating range of the instrument. Equating these quantities and dividing by J_s there results the differential equation of motion of the ring

$$\frac{d^2\theta}{dt^2} + \frac{I^2m^2\theta}{J_sL_s} = 0,$$

which is recognized as the equation of simple harmonic torsional oscillation of frequency F, where

$$4\pi^2F^2 = I^2m^2/J_sL_s, \text{ from which}$$

$$I = \frac{2\pi F}{m} \sqrt{J_sL_s}\, F, \text{ in e.m.u.}$$

Thus the effective value of current measured is independent of wave form and frequency above one megacycle, where the instrument would be used, and is directly proportional to the frequency of torsional oscillation of the ring.

$$I = CF \text{ amperes,} \qquad \text{where} \qquad C = \frac{2\pi\sqrt{J_s L_s}}{m}.$$

The instrument constant C may be obtained experimentally by comparison with a thermocouple instrument at a frequency of the order of one to five megacycles where the thermoammeter is accurate, or directly by calculation from physical data of the electrodynamic ammeter. A comparison of these methods is shown by Fig. 3, the points were obtained experimentally and the continuous curve was obtained by calculation. This check is satisfactory considering that F, the mechanical frequency of oscillation, was determined by means of a stop watch, the readings of which were reproducible to within ½ per cent.

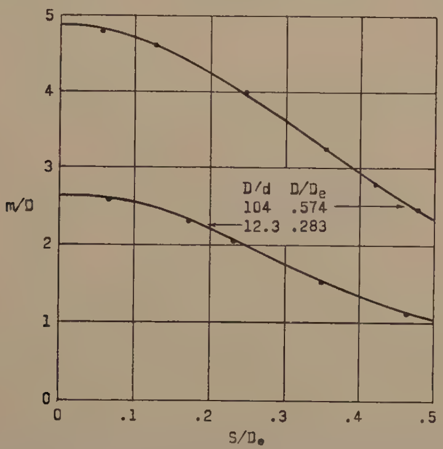


Fig. 3—D is the diameter of the ring; d, the diameter of the exciter coil; and S, the distance from the plane of the exciter to the center of the ring.

Experimental Results

There is good agreement between the electrodynamic ammeter and the five-ampere thermocouple type up to about five megacycles, and fairly good up to ten megacycles, above which the thermo instrument reads progressively higher as indicated in the following table and plotted in Fig. 4:

10 Mc	4%
20	13
40	32
60	54
80	80

A considerable portion of this difference may be attributed to skin effect in the thermoelement; a calculated curve is shown in Fig. 4. For

thermo instruments of lower range this difference would be less and of higher range, greater. At higher frequencies the effect of parallel resonance in the thermo instrument appears and as the element is in one of the parallel branches a larger current will flow through it than in the line leading to the instrument, which is the current desired. The problem is still further complicated by capacitance to the case, impedance in the ground lead to the case and other circuit configuration difficulties.

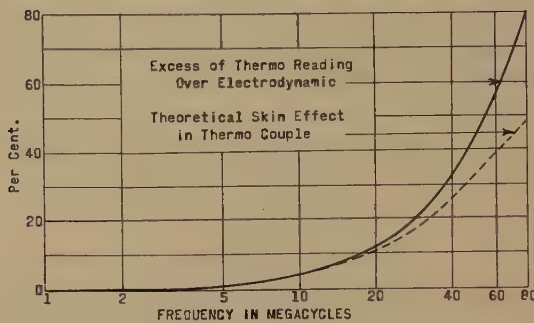


Fig. 4—Comparison of experimental results.

When using the electrodynamic ammeter for measuring currents of such low magnitude that the mechanical frequency of oscillation of the secondary ring is only slightly greater than its natural frequency a correction factor should be used. Where it is twice the natural frequency the correction factor is 0.87 and where it is five times the correction factor is 0.98.

It is remarkable that two instruments differing so widely in principle of operation would agree so completely when operating at frequencies up to ten megacycles. This gives added faith in both methods.

The object of this investigation, which was to develop a means for checking thermoammeters up to 100 megacycles or higher, has been partially achieved, but the study should be continued at still higher frequencies.

Bibliography

(1) E. B. Moullin, "The development of a precision ammeter for very high frequencies," *Jour. I.E.E.* (London), vol. 68, p. 544, (1930).

(2) C. L. Fortesque and L. A. Moxon, "An ammeter for very high frequencies," *Jour. I.E.E.* (London), vol. 68, p. 556, (1930).

(3) C. L. Fortesque, "The theory and design of hot-wire ammeters for frequencies of 25 to 100 megacycles," *Jour. I.E.E.* (London), vol. 79, p. 179, (1936).

(4) J. H. Miller, "Thermocouple ammeters for ultra-high frequencies," Proc. I.R.E., vol. 24, pp. 1567–1572; December, (1936).

(5) J. D. Wallace and A. H. Moore, "Frequency errors in radio-frequency ammeters," Proc. I.R.E., vol. 25, pp. 327–339; March, (1937).

Proceedings of the Institute of Radio Engineers
Volume 25, Number 11 *November, 1937*

SOME NOTES ON RAIN STATIC IN JAPAN*

By
Tomozo Nakai
(Tohoku Imperial University, Sendai, Japan)

THE paper on "Rain Static" by H. K. Morgan[1] showed that rain or snow sometimes disturbs aircraft radio reception and presents an important problem in aircraft radio communication.

Land as well as aircraft radio reception is also disturbed by rain but only rarely. A serious disturbance of this type was experienced in 1929 at the Isohama Branch of the Electrotechnical Laboratory of the Ministry of Communication of Japan. In the "Report of the Radio Division" of the laboratory the author has pointed out that such a disturbance is caused by rain and some theoretical calculations with respect to this problem were given.[2] The relation of static to meteorological conditions at that time was as follows:

At 1030 to 1200, 1300 to 1430, and 1500 J.C.S.T. on September 13, 1929, very strong static of very long duration such as grinders was received on a T antenna which had been tuned to about twenty kilocycles. The characteristics of the grinders were as follows:

(a) The strength of the grinders was nearly 1000 to 1700 microvolts per meter (normal strength at this frequency in this season is about 30 to 40 microvolts per meter) and the signal of HZA (Saigon) which has a strength of about 400 microvolts per meter was completely masked by the grinders.
(b) The grinders were not experienced on an indoor loop antenna.
(c) The grinders started with rain and faded away when the rain stopped.
(d) Whenever thunder was heard the grinders faded away for several seconds and during this period the signal of HZA was received normally.

A downpour of heavy rain occurred at the time stated above and at 1040 to 1230 J.C.S.T. weak thunder was heard at Mito which is about 10 kilometers away from Isohama; at 1200 to 1330 J.C.S.T. there was strong thunder at Fukushima about 100 kilometers away, and at 1530 to 1615 J.C.S.T. weak thunder at the lower part of the

* Decimal classification: R114. Original manuscript received by the Institute, May 4, 1937. Submitted originally as a discussion to H. K. Morgan's paper, "Rain Static[1]."
[1] Proc. I.R.E., vol. 24, pp. 959–963; July, (1936).
[2] Tomozo Nakai, October, (1929).

Sagami River about 130 kilometers away from Isohama. Each of these thunderstorms covered only a small range. The weather chart at 600 J.C.S.T. on this day is shown in Fig. 1, and at 1800 J.C.S.T. it was nearly the same as at 600.

Although thunder was heard on this day the grinders were not due to lightning discharges. The sounds of the static were quite different from static caused by lightning discharge, and the time of the grinders

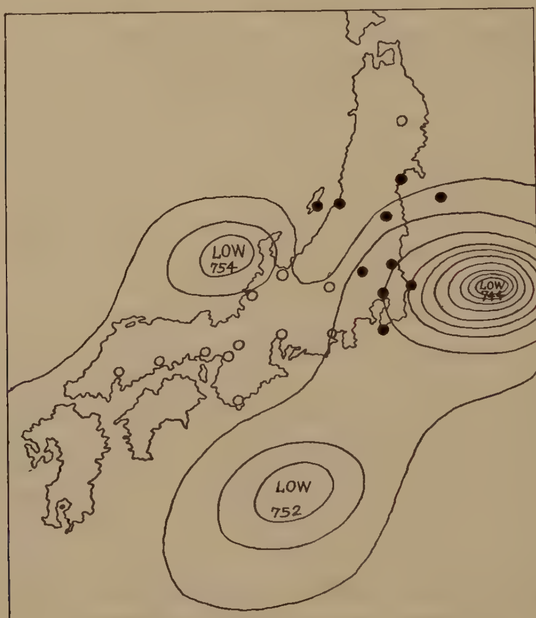


Fig. 1

did not correspond with the lightning discharges. A study of the characteristics of these grinders suggests that the electric charge of rain is the cause of the disturbance. Data on the electric charge of rain and the dimensions of a raindrop on this day are not available, but according to the record of the past,[3] the electric charge of a raindrop is ordinarily considered as about 0.4 to 0.5 electrostatic units per cubic centimeter and in the extraordinary case it is as large as 40 electrostatic units per cubic centimeter. A raindrop is generally considered as having a diameter up to 0.3 millimeter for a small drop, up to 3 millimeters for a medium drop, and from 5 to 6 millimeters for a large drop.

If we assume the diameter of a raindrop on this day to have been

[3] T. Okada, "Meteorology," (in Japanese), p. 312 and p. 421, (1927).

3 millimeters and the electric charge as 40 electrostatic units per cubic centimeter, then the charge of a raindrop on this day was 0.566 electrostatic units.

The hourly rainfall on this day observed by the Mito Meteorological Station was as follows:

TABLE I

J.C.S.T.	millimeters of rain
500	5.0
600	5.8
700	7.7
800	13.5
900	8.9
1000	14.4
1100	2.5
1200	2.8
1300	0.5
1400	0.1
1500	—
1600	0

An average rainfall of about 8 millimeters per hour was observed which is equivalent to 0.8/3600 cubic centimeters per square centimeter per second and is also equivalent to about 0.016 drop per centimeter per second.

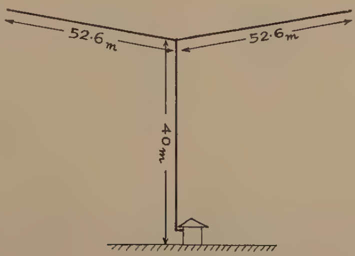


Fig. 2

The antenna used for the station is shown in Fig. 2, and is of seven-strand wire each strand being 1.626 millimeter in diameter. If the rain struck along a plane perpendicular to the T antenna at an angle of about 60 degrees to the ground the cross-sectional area of the antenna becomes about 6110 square centimeters. That is, about 98 raindrops struck the antenna per second.

The fundamental equation for the discharge current of an antenna is

$$L\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{1}{K}i = 0$$

$$i = \frac{dq}{dt},$$

where,

L = total inductance of the antenna,
R = total effective resistance of the antenna,
K = total capacitance of the antenna, and
q = charge of the antenna.

If $R^2 \ll \frac{4L}{K}$, as it is in this case, we have

$$i = -\frac{Q_0}{nKL}\epsilon^{-(R/2L)t}\sin nt,$$

where,

and
$$n = \sqrt{\frac{1}{KL} - \frac{R^2}{4L^2}}$$

$$Q_0 = \text{charge of } K \text{ at } t = 0.$$

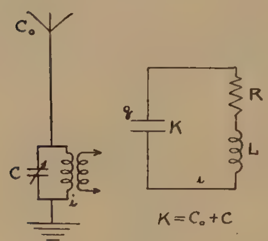


Fig. 3

In substituting the numerical values of 120 millihenrys (1.33×10^{-13} electrostatic units), 1000 ohms (1.11×10^{-9} electrostatic units), and 0.00053 microfarad (475 electrostatic units) for L, R, and K, respectively, we have

$$n = 12.55 \times 10^4$$

and
$$\frac{R}{2L} = 4160.$$

Putting $Q_0 = 0.566$, we have

$$\frac{Q_0}{nKL} = 7.14 \times 10^4 \text{ electrostatic units}$$

$$= 23.8 \text{ microamperes.}$$

The numerical values of I_1, I_2, I_3, etc. (Fig. 4) are given in Table II. Thus, the charge of one drop gives a damped antenna current as

shown in Fig. 4, and if we assume that 100 drops struck the antenna per second at regular intervals of 0.01 second, such damped currents would flow in the antenna at regular intervals of 0.01 second (Fig. 5). It can reasonably be appreciated that such a group of damped waves may disturb reception as a form of grinders.

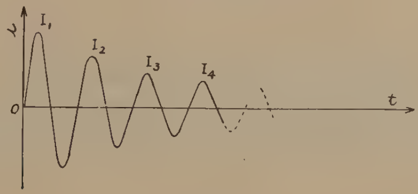


Fig. 4

The received current of 20 microamperes in this antenna corresponds to a field strength[4] of 1400 microvolts per meter at $f=20$ kilocycles. The maximum value of the current caused by the discharge

TABLE II

I	microamperes
I_1	22.7
I_2	18.3
I_3	15.0
.	.
.	.
.	.
I_{10}	3.5
.	.
.	.
.	.

of one drop on this day corresponds to a field strength of about 1400 microvolts per meter. The strength of the grinders on this day was 1000 to 1700 microvolts per meter as stated above, and this value

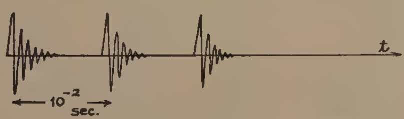


Fig. 5

corresponds to 5000 to 8500 microvolts per meter absolute.[5] Thus the measured value is of the same order as that calculated theoretically. It may be concluded that rain may disturb radio reception and the strong grinders on September 13, 1929, may reasonably be interpreted

[4] Researches of the Electrotechnical Laboratory, Ministry of Communications of Japan, No. 217.

[5] The strength of grinders was measured in relative units. The unit corresponds to about one-fifth the absolute field strength.

as rain static. In case of an airplane the dimension of the antenna used is generally smaller than that on the land station but an airplane moves with a high velocity. This is equivalent to an enlargement of the antenna in that the number of drops of rain which strike it per second is large though the dimension of the antenna is small. Consequently, an audible noise may result.

Acknowledgment

This paper is published by permission of Dr. Y. Kusunose, chief of the radio division of the Electrotechnical Laboratory, Ministry of Communications of Japan, and the author thanks him for his courtesy.

Proceedings of the Institute of Radio Engineers
Volume 25, Number 11 *November, 1937*

A THERMAL METHOD FOR MEASURING EFFICIENCIES AT ULTRA-HIGH FREQUENCIES APPLIED TO THE MAGNETRON OSCILLATOR*

By
H. W. Kohler
(Cruft Laboratory, Harvard University, Cambridge, Mass.)

Summary*—Efficiency measurements at ultra-high frequencies are briefly discussed. A method which uses a thermocouple located within the tube envelope as an indicator is described, and theory and results for the magnetron oscillator of dynatron type are given.*

Introduction

At ultra-high frequencies, efficiency measurements generally are made by determining either tube loss or output directly, instead of making voltage and current measurements. The photometric method of efficiency measurement is the one most widely used. An incandescent lamp is connected as a load to the tank circuit and dissipates the high-frequency power. The brilliancy of the lamp is measured, and then the direct-current or low-frequency power giving the same brilliancy is determined, and thus the available power output is obtained. Efficiencies based on this method are slightly low as they take account of the power loss between generator and load. Calibration requires handling of the load lamp which should have a thin straight filament.

In the thermal method the heat radiation from the tube elements is measured under oscillatory and nonoscillatory conditions, whereby in the latter state the input power is so adjusted as to give the same reading of the thermal indicator as in the oscillatory state. If the filament power is the same for both measurements the difference in plate power input in the two cases is the available power output, and thus the efficiency is found.

The New Thermal Method

In the new method the thermal indicator is a thermocouple located within the tube envelope. This makes the indicator independent of drafts and ordinary changes in room temperature. Calibration does not require manipulation of the load and the measuring apparatus is very simple.

The thermocouple consists of tungsten-constantan wires five mils in diameter. The junction is located about $\frac{3}{8}$ inch from the plate as

* Decimal classification: R290×R339. Original manuscript received by the Institute, January 21, 1937; revised manuscript received by the Institute, May 5, 1937; July 22, 1937.

shown in Fig. 1. Fig. 2 shows the magnetron tube used. A by-pass condenser is put across the thermocouple terminals at the tube socket,

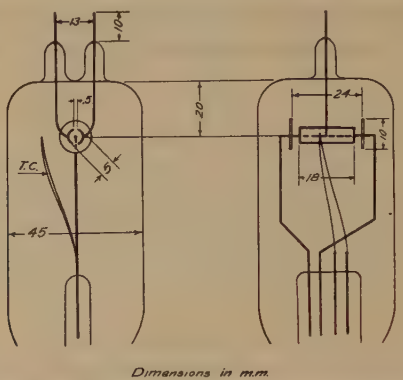


Fig. 1—Construction details of magnetron with thermocouple.

and the connection to the microammeter is made with thin twisted resistance wires. Removal of shielding of meter and of thermocouple leads gave a smaller thermal current when the tube was in the oscillatory state.

Fig. 2—Close-up of magnetron.

Efficiencies obtained checked figures published by other workers in this field. Photometric measurements of the light intensity of the lamp loading the tank circuit were compared with those made with direct-current operation of the lamp; the efficiency obtained by the thermo-

couple and the photometric method agreed well, the latter giving somewhat lower efficiency. From the nature of the thermal method of efficiency measurement it follows that any residual induced currents in the thermocouple circuit will give too low an efficiency; i.e., make the results conservative.

The time between successive equilibriums of temperature is one to three minutes. This presumes that the cold tube has been heated up for about fifteen minutes and that the supply voltages are constant. The electromotive force developed by the thermocouple gives readings up to ten or more microamperes, and the use of a bridge circuit is unnecessary.

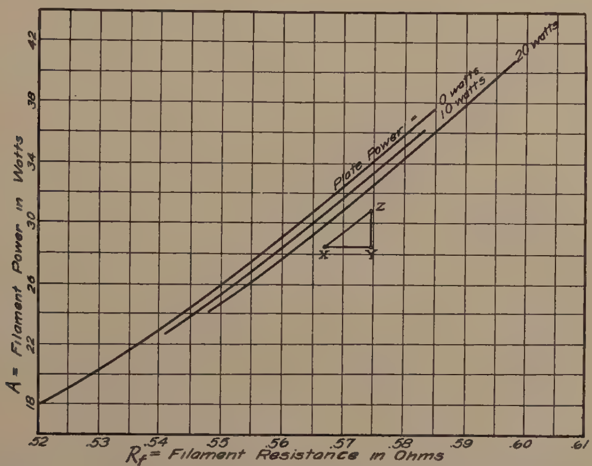


Fig. 3—Filament resistance versus filament power, for plate power equal to 0, 10, 20 watts. (Nonoscillatory state.)

As the magnetron is brought from the nonoscillatory to the oscillatory state a decrease in filament current and an increase in filament voltage; i.e., an increase in filament direct-current resistance is observed. This increase in filament resistance is caused by electrons returning to the filament and hitting it with velocities greater than the initial velocity of emission. Thus a transfer of plate battery power ΔB into the filament circuit takes place. Presumably secondary emission is produced at the filament.

In analytical form we may write for the oscillatory state

$$D = a(A + \Delta B) + b(B_1 - \Delta B - B_0). \tag{1}$$

For the nonoscillatory state we have

$$D = aA + bB_2 \tag{2}$$

where,

D = thermocouple reading,

A = filament power,
B_1, B_2 = plate power input,
B_0 = plate power output.
a and b are factors of proportionality (microamperes/watt) and are both functions of A and B.

By operating adjustments D and A in (1) and (2) are held constant, and from experimental evidence shown in Fig. 4 the quantities a and b may be considered constant in the above two equations.

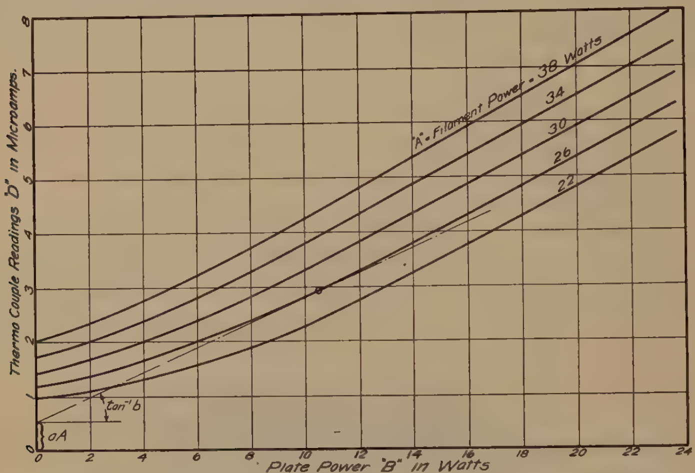


Fig. 4—Thermocouple readings versus plate power; filament power as parameter.

Solving for the unknown plate power output B_0 we have

$$B_0 = B_1 - B_2 - \Delta B\left(1 - \frac{a}{b}\right). \tag{3}$$

The plate efficiency of the tube is defined as

$$Y = \frac{B_0}{B_1} = 1 - \frac{B_2}{B_1} - \frac{\Delta B}{B_1}\left(1 - \frac{a}{b}\right). \tag{4}$$

The first two terms give the efficiency as ordinarily used in the thermal efficiency measurements. The third term gives the correction due to transferred plate power ΔB. It is always negative since $a < b$.

In absence of a better method the transferred plate power ΔB is found as follows:

In Fig. 3 filament resistance R_f, is plotted against filament power A. The increment in filament resistance between the oscillatory and non-oscillatory state is $(R_{f_0} - R_{fn})$ and corresponds to an increment in filament power ΔA, which is equal to the transferred plate power ΔB.

$$\Delta B = (R_{f_0} - R_{f_n}) \frac{dA}{dR_f} = \overline{YZ}.$$

$$R_{f_0} - R_{f_n} = \overline{XY}.$$

The plot of ΔB versus frequency is approximately a straight line. ΔB increases with frequency and is greatest at the shortest wave length.

The factors a and b are determined from a plot of thermocouple readings versus plate power, with the filament power as parameter, taken under nonoscillatory conditions. (Fig. 4.) Considering the equation

$$D = aA + bB$$

it is seen that the slope of the curves equals b, and the intercept at the axis of ordinates equals aA. Determination of a and b should be made from the operating point under oscillatory conditions, for filament power $(A+\Delta B)$ and plate loss $(B_1-B_0-\Delta B)$. However, B_0 is not known. Since the curves in Fig. 4 are practically straight parallel lines for plate losses over ten watts, B_2 may be used as the plate loss in finding a and b.

The following sources of error may be present in the new method:

The thermocouple may be heated by induced currents (including eddy currents) and displacement currents.

Induced voltages in the thermocouple itself are made small by having only its junction enter the radio-frequency region of the tube and by placing the plane of its closely spaced wires at right angles to the plane of the two plate leads. Eddy currents are small because of the thin wires of the couple and the resistive material used. Displacement currents are negligible.

A further error may be introduced if a given plate loss does not give the same thermal current reading in the oscillatory and nonoscillatory states due to a different heat pattern at the plate in the two cases. This error seems to be small as the plates appeared uniformly heated, a condition which was aided by the rather heavy plate materials (ten mils). Also, the distance from thermal junction to plate is relatively large (about $\frac{3}{8}$ inch) so that the effect of any plate regions of unequal temperature is averaged out.

Errors due to changes in magnetizing coil temperature are very small because the coils were prevented from cooling off in the nonoscillatory periods as then they were switched into series-opposing connection.

An error in values of efficiency, obtained where the correction due to transferred plate power ΔB is made, may be present if the power transfer is not uniform along the filament. This, however, is not an error in the thermal method so much as an error in determining the correction due to ΔB.

If both A and ΔB are uniformly distributed along the filament, the direct-current resistance of the filament is constant if $(A+\Delta B)$ and the plate power are constant, no matter in what proportion A and ΔB are. Actually the transferred power tended to be more concentrated into the central part of the filament as the frequency was increased. The resistance measured became too large with increasing frequency, and the corrected efficiency obtained too small.

In Fig. 5 maximum efficiencies are given as a function of the wave length. The uncorrected curve is directly obtained from the thermal indicator readings which can be reproduced within ± 2 per cent by successive readings. The accuracy is mainly limited by fluctuating plate

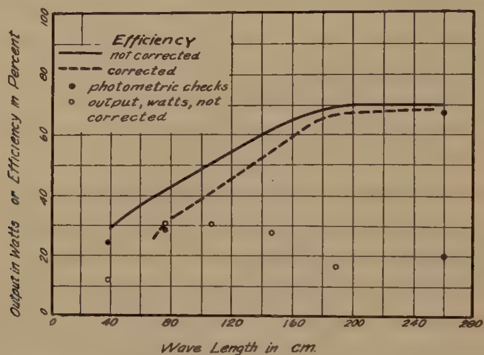


Fig. 5—Maximum efficiencies measured by thermal method.

and filament voltage. The dashed curve takes account of the power transfer ΔB and is believed to be correct within ± 5 per cent. It is limited by the accuracy with which the transferred power can be measured.

At the shortest wave length obtainable for dynatron type oscillations the corrected efficiency becomes practically zero. This is because the calculated transferred power is too large. From the brilliance of the lamp loading the circuit it could be seen that an appreciable radio-frequency power output was present. The filament of the lamp which had the shape of a polygon was not uniformly lighted over its entire length at the highest frequency, probably because of capacitive effects inside the bulb between supporting wires. Therefore, the photometric check of this point of the curve is not reliable.

Acknowledgment

The writer wishes to express his sincere appreciation to Professor E. L. Chaffee of Harvard University for his interest and valuable discussions.

Proceedings of the Institute of Radio Engineers
Volume 25, Number 11 *November, 1937*

A LOW DISTORTION AUDIO-FREQUENCY OSCILLATOR*

By
Herbert J. Reich
(Department of Electrical Engineering, University of Illinois, Urbana, Illinois)

Summary—*In the first portion of this paper, which deals with the theory of negative resistance oscillators, it is shown that for small harmonic content the form of the negative resistance characteristic in the vicinity of the operating point should be such that the average negative resistance increases with amplitude of oscillation. It is then shown that Turner's Kallirotron circuit has negative resistance characteristics of the desired form.*

The final part of the paper describes an audio-frequency oscillator based on Turner's circuit. The second and third harmonic content of the output may be kept below 0.2 per cent by the use of low resistance coils and diode automatic amplitude control. Higher harmonics are negligible. At a thousand cycles per second the frequency drift relative to a tuning fork oscillator does not exceed 0.04 cycle. The frequency change caused by a 22½-volt change in plate supply voltage is 0.04 cycle.

THE good frequency stability and low harmonic content of the dynatron oscillator are offset by the dependence of the action of the dynatron upon secondary emission.[1] Large variations in secondary emission during the life of the tube, and differences in the secondary emission characteristics in individual tubes of the same type, seriously affect the operation and calibration of dynatron oscillators. Several investigators have suggested the desirability of the development of special tubes which exhibit more stable secondary emission characteristics, but so far special dynatron tubes have not been made available. The improved form of the van der Pol oscillator recently described by Herold[2] represents one effort to develop an oscillator having the advantages of the dynatron without its disadvantages. No study appears to have been made of the suitability of the "Kallirotron" negative resistance circuit of Turner[3] as the basis of a negative resistance oscillator. It is the purpose of this paper to show that the characteristics of this circuit are ideal for its use in an oscillator and to describe a low distortion oscillator based upon it.

Negative Resistance Oscillator Theory

The basic circuit of a negative resistance oscillator is shown in Fig. 1, in which the dotted element represents the resistance of any device

* Decimal classification: R355.9. Original manuscript received by the Institute, May 17, 1937.

[1] M. G. Scroggie, *Wireless Eng. and Exp. Wireless*, vol. 10, p. 527, (1933). Includes a bibliography of thirty-five items on dynatron oscillators.

[2] E. W. Herold, Proc. I.R.E., vol. 23, pp. 1201–1223; October, (1935). Includes a bibliography of fifty-five items on negative resistance.

[3] L. B. Turner, *Radio Rev.*, vol. 1, p. 317, (1920).

which has a negative resistance; i.e., one whose current-voltage characteristic has a negative slope. Although the negative resistance, which is the reciprocal of the slope of the current-voltage curve, is not in general constant, it may be assumed to be constant over a small range of current. Under this assumption the application of Kirchhoff's laws

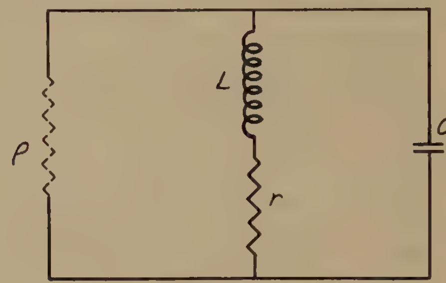


Fig. 1

to the circuit of Fig. 1, and solution of the resulting differential equation gives the following equation for the current in the circuit:

$$I = A\epsilon^{-(r/L+1/\rho c)t} \cos \omega t, \tag{1}$$

in which A is a constant, and the radian frequency of oscillation is

$$\omega = \sqrt{\frac{r+\rho}{\rho}\frac{1}{LC} - \frac{1}{4}\left(\frac{1}{\rho C} + \frac{r}{L}\right)^2}. \tag{2}$$

Since the numerical value of ρ is negative, it can be seen that if the magnitude of ρ is greater than L/rC, the amplitude decreases with time, and oscillation finally ceases; if the magnitude of ρ is less than L/rC, on the other hand, the amplitude builds up. In the critical case in which ρ is equal to L/rC, the exponential factor is unity, indicating that oscillation neither builds up nor dies down, but, once started, continues with constant amplitude. The criterion for oscillation is

$$|\rho| \leqq \frac{L}{rC}. \tag{3}$$

Under the threshold condition, when $|\rho| = L/rC$, the frequency of oscillation is

$$f = \frac{1}{2\pi}\sqrt{\frac{r+\rho}{\rho}\frac{1}{LC}}. \tag{4}$$

Solution of the circuit equations under the more general assumption that ρ is not constant leads to the expected result that curvature of the current-voltage characteristic introduces harmonics into the current. Curvature of the characteristic also causes $\bar{\rho}$, the average negative resistance during oscillation, to differ from the initial value of ρ at the operating point. The variation of $\bar{\rho}$ with amplitude of oscil-

lation explains the limitation of oscillation amplitude. Equilibrium amplitude is dependent not only upon the circuit parameters, but also upon the shape of the current-voltage characteristic. If the characteristic is of the form shown in Fig. 2a, and the operating point is at O, increase of amplitude will first cause a decrease in $\bar{\rho}$, and hence further increase of amplitude. When the amplitude becomes sufficiently great to extend the path of operation beyond the points of inflection at a, then $\bar{\rho}$ will begin to increase with further increase of amplitude and will finally become sufficiently high so that the critical condition for

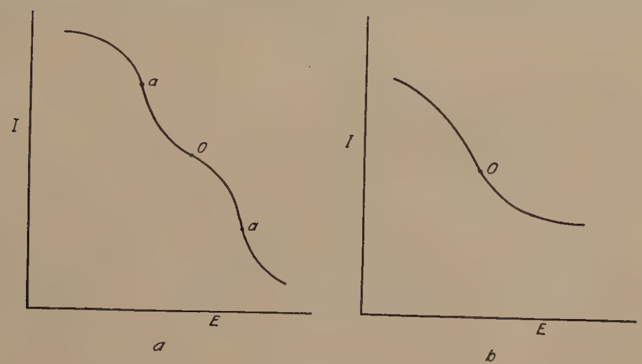


Fig. 2

oscillation is reached. Any further increase of amplitude would make $\bar{\rho}$ large enough so that the exponent of (1) would be negative, and the amplitude would fall. Hence, equilibrium is established at such an amplitude that the criterion for oscillation is just satisfied. Because the equilibrium amplitude may be large, and the path of operation is curved, the harmonic content may be excessive. If, on the other hand, the characteristic curve and operating point are as shown in Fig. 2b, then $\bar{\rho}$ increases continuously with increase of amplitude, and the amplitude of oscillation and harmonic content may be made as small as desired by making the value of ρ at the operating point as nearly as possible equal to L/rC. The fact that the characteristic curves of dynatrons approximate the form of Fig. 2b accounts for the low harmonic content attainable with dynatron oscillators.

Kallirotron Circuit

The "Kallirotron" circuit is shown in basic form in Fig. 3. By the application of the equivalent plate circuit theorem to this circuit it may be shown that the resistance between points A and B has the value[4]

[4] J. A. Stewart, bachelor's thesis: "The Roberts neutralizer circuit," Purdue University, June, (1935).

$$\rho = \frac{2r_p}{\dfrac{r_p + R_b}{R_b} - \mu}. \tag{5}$$

ρ is negative if μ is greater than $(r_p+R_b)/R_b$. A somewhat simpler analysis, which gives an insight into the physical behavior of the circuit is given in the Appendix. The current-voltage characteristics for the circuit of Fig. 3 were obtained by measuring the direct current, I, flowing into B as the result of the application of a voltage, E, between A and B. Figs. 4 and 5, obtained with a type 53 twin triode,

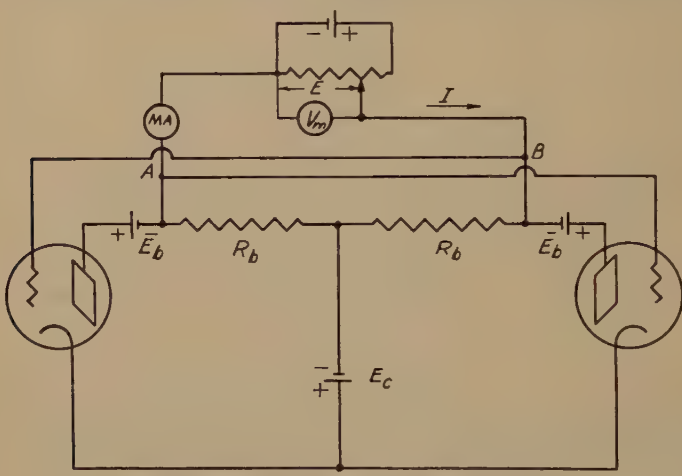


Fig. 3—Basic negative resistance circuit.

show the manner in which the characteristics are affected by grid and plate supply voltages, E_c and E_b. Fig. 6 shows how the negative resistance, found from the reciprocal of the slope at the origin, varies with grid and plate supply voltages, and with plate load resistance. These curves were found to be in excellent agreement with similar curves determined[5] from (5).

Since the resistance between A and B of the circuit of Fig. 3 may be made negative, a negative resistance oscillator can be made by connecting a parallel resonant circuit between points A and B. Furthermore, since the characteristic curves are of the desirable form of Fig. 2b, the amplitude and harmonic content may be made as small as desired by adjusting the circuit parameters or supply voltages. In Fig. 7 is shown a practical form of oscillator which uses a common plate

[5] The change in r_p resulting from the change of operating voltage with R_b must be taken into account in comparing theoretical and experimental curves of ρ against R_b.

supply voltage for both tubes. C_c and R_c should be sufficiently large so that the reactance of C_c is small compared to R_c at the lowest oscil-

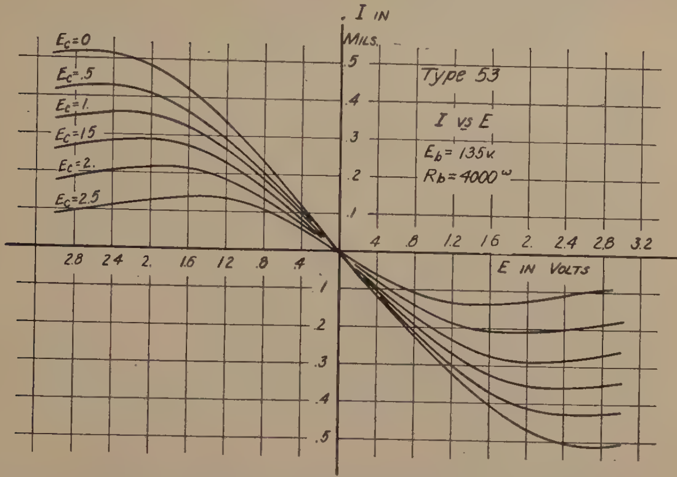


Fig. 4—Negative resistance characteristics for the circuit of Fig. 3, using type 53 twin triodes.

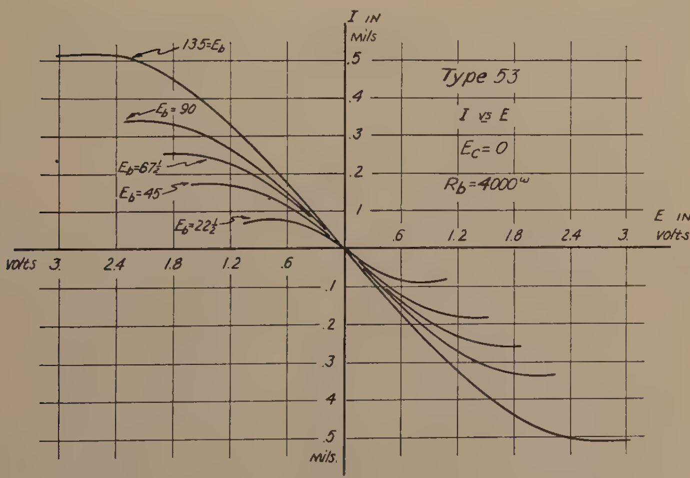


Fig. 5—Negative resistance characteristics for the circuit of Fig. 3, using type 53 twin triodes.

lation frequency. Because of the low reactance of C_c, the tuned circuit may be connected between the grids, as shown by the dotted lines in Fig. 7, instead of between the plates. Fig. 8 shows another form of circuit in which the coupling condensers, C_c are replaced by resistors, R_c'.

Amplitude Control

In order to keep the amplitude of oscillation, and hence the harmonic content, small, it is necessary to adjust the circuit parameters or supply voltages so that L/rC is only slightly larger than ρ at the operat-

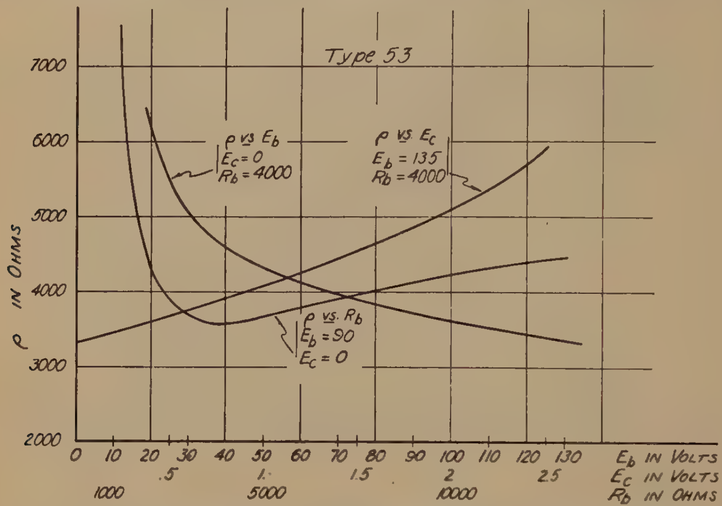


Fig. 6—Variation of negative resistance with supply voltages and load resistance; type 53 twin triode.

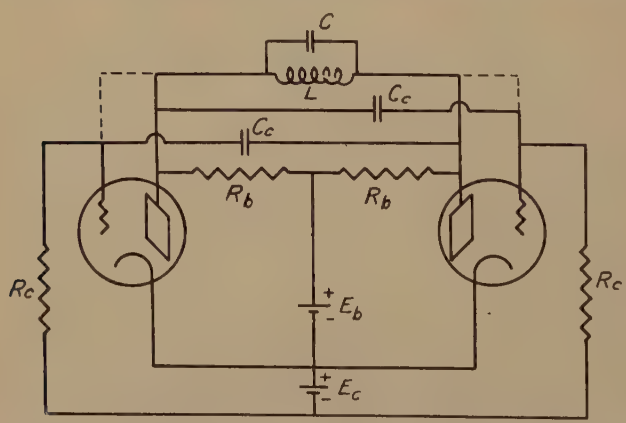


Fig. 7—Balanced negative resistance oscillator with common B voltage.

ing point. As the frequency is varied by changing L or C, it is therefore also necessary to change ρ or r. r may be varied by means of a variable resistance in series with the inductance, but it has been found that an increase of r raises the harmonic content. ρ can be varied by means of

the supply voltages or R_b, as indicated by the curves of Fig. 6. ρ can also be varied if the resistors R_b are replaced by potentiometers, as shown in Fig. 10. If the oscillator is to be calibrated at selected frequencies, then the potentiometer settings can be included in the calibration chart. A more satisfactory method is to use automatic amplitude control. A certain amount of automatic control is obtained as the result of flow of grid current in the circuit of Fig. 7 when E_c is small. When the amplitude builds up to the point at which grid current starts flowing, the grid sides of the condensers C_c accumulate a negative charge

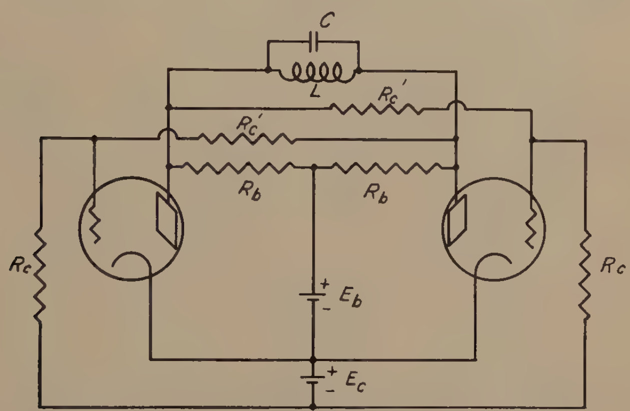


Fig. 8—Balanced negative resistance oscillator with common B voltage.

faster than it leaks off through the resistors R_c. Consequently the grids become more negative, and $\bar{\rho}$ is increased, finally becoming high enough to establish equilibrium. Much better control is obtained by rectifying the oscillator voltage by means of a diode and using the rectified voltage as grid bias for the oscillator.[6] The sensitivity is increased by the use of a 3:1 audio transformer to step up the voltage before rectification. A circuit which incorporates this type of amplitude control is shown in Fig. 9. Although this method does not hold the output level strictly constant, the variation is sufficiently small so that the harmonic content may be kept at a low value over the whole audio-frequency range.

Diode amplitude control does not prevent change of amplitude with plate supply voltage. For given values of r, L, and C, the criterion for oscillation is satisfied for a particular value of ρ. Fig. 6 shows that ρ decreases with increase of E_b. The amplitude will therefore rise until the rectified bias voltage increases sufficiently to bring $\bar{\rho}$ to the critical value. The variation of amplitude with plate supply voltage may be

[6] L. B. Arguimbau, Proc. I.R.E., vol. 21, pp. 14–28; January, (1933); J. Groszkowski, Proc. I.R.E., vol. 22, pp. 145–151; February, (1934).

reduced by increasing the transformation ratio of the transformer which feeds the diode circuit. Similar reasoning shows that the amplitude of oscillation may be adjusted by means of a cathode resistor in the oscillator circuit. If the bias produced by the cathode resistor is changed by an amount ΔE_c, the crest oscillator voltage will change by $\Delta E_c/n$, where n is the transformation ratio of the transformer.

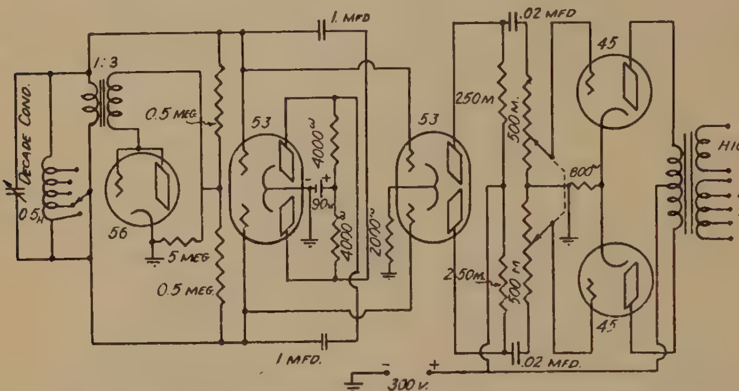


Fig. 9—Negative resistance oscillator with diode amplitude control.

Effect of Coil Design on Harmonic Content

Tests were made with a General Radio wave analyzer to determine the effect of coil design upon harmonic content. It was found that the harmonic content increases with the resistance of the inductance coil and with resistance introduced in series with the coil or between the tuned circuit and the plates. Examination of the dynamic negative resistance characteristic by means of a cathode-ray oscillograph indicated that the increase of harmonic content results from a tendency of the circuit to unbalance because of slight differences in the two triodes. The unbalance can be partially corrected by individual adjustments in the resistors R_b. The harmonic content increases rapidly as the coil or circuit resistance approaches the magnitude of the negative resistance, and eventually the oscillation changes into the relaxation type. This is to be expected, since the circuit becomes that of a multivibrator if the inductance is omitted.

Using a tapped inductance, it was found that the harmonic content at a given frequency was always least when the tuned circuit contained the smallest part of the inductance with which oscillation of given amplitude could be obtained. A slight reduction of harmonic content results from short-circuiting the unused portion of the inductance. When the inductance exceeds four or five henrys a portion of

the distortion results from the iron core of the amplitude control transformer, the primary of which shunts the tuned circuit.

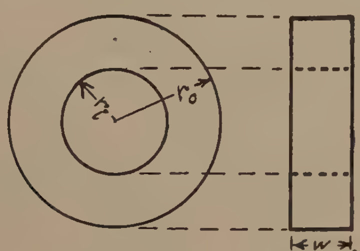
A second advantage of low coil resistance results from the fact that the frequency stability increases with decrease of resistance. Equation (4) shows that the dependence of frequency upon ρ decreases with r. If r is small in comparison with ρ, changes in ρ resulting from changes of battery voltages will have little effect upon frequency.

The following table gives the design specification of three coils,

TABLE I

Coil Design Data

Coil No.	L Henrys	R Ohms	DCC Wire No.	Wt. Lbs.	Turns	Layers	Taps	Dimensions Inches			% H_2	% H_3	Min. freq. c.p.s.
								r_i	r_o	w			
1	11.5	2100	32	3	10,700	110	11	1.6	3.2	1.9	0.5	0.5	40
2	2.8	280	26	5	5,000	66	11	1.6	3.2	1.9	0.25	0.25	100
3	0.48	23	18	15	2,200	44	11	2.0	2.0	2.4	0.18	0.18	120



together with the average harmonic content, and the minimum frequency at which oscillations can be obtained with a plate supply voltage of 90 volts. All coils are designed for optimum L/R ratio.[7]

The fourth and higher harmonics are negligible. It is evident from this table that the harmonic content can be reduced to a very low value by the use of large wire, but that reduction of harmonic content is accompanied by a marked increase of weight if the inductance is maintained sufficiently high to insure low minimum frequency. For most applications coil No. 2 represents a satisfactory compromise between low weight on one hand, and small harmonic content and high stability on the other. The minimum frequency can be reduced with comparatively little increase of harmonic content by raising the inductance of coil No. 2 by increasing the amount of wire. Tests are planned to determine the practicability of using coils with low-loss iron cores.

[7] Morgan Brooks and H. M. Turner, "Inductance of coils," Bulletin No. 53, University of Illinois Engineering Experiment Station.

Complete Oscillator

Fig. 9 shows the circuit diagram of the complete oscillator. Type 53 triodes are used in the oscillator stage and in a stage of voltage amplification which feeds the final power stage of 45's in push-pull. Any type of tube may be used as the amplitude control diode. With coil No. 3 the audible frequency range can be adequately covered by means of a decade arrangement of condensers in four banks of 0.001–0.009, 0.01–0.09, 0.1–0.9, and 1–4 microfarads capacitance. The 1–4-microfarad bank may be omitted with coils No. 1 and No. 2.

In the original design of the oscillator the plate voltage of the oscillator stage was obtained from the same power supply as that of the amplifier stages. It was found, however, that fluctuations of line voltage produced considerable variation of output level. This resulted not only from the direct effect of changes of plate voltage of the oscillator stage, but also from associated changes of grid voltage caused by the plate-to-grid coupling. The difficulty can be remedied by the use of a regulated B supply, but since the oscillator triodes draw a total current of only about 4 milliamperes, it is simpler to use two small 45-volt B batteries to supply the oscillator voltage.

The output level is adjusted by means of twin potentiometers which control the input to the power stage. The maximum power output is approximately sixty milliwatts. Because it is operated so far below its normal full output, the amplifier makes a negligible contribution to the harmonic content of the output. When the amplitude control tube is removed, the harmonic content rises to about two per cent, most of which is accounted for by the third harmonic. The power output may be increased to about 200 milliwatts by the use of a 79 tube in place of the 53 in the first amplifier stage.

Frequency Stability

To obtain a conservative measurement of frequency stability, readings were made with coil No. 1, which has the highest resistance. The frequency was adjusted to 1000 cycles, using two sections of the coil. The resistance of this portion of the coil was 320 ohms. Reduction of plate supply voltage from 90 volts to $62\frac{1}{2}$ volts resulted in a frequency change of 0.04 cycle per second. The maximum frequency drift relative to a 1000-cycle tuning fork oscillator from the time of commencement of oscillation was also approximately 0.04 cycle per second. Higher stability was obtained with coils No. 2 and No. 3. With coil No. 3 the frequency change from 1000 cycles resulting from a $22\frac{1}{2}$-volt change in plate supply voltage was less than 0.03 cycle per second. In determining the frequency stability, no attempt was

made to control or vary the temperature. Since the circuit elements are not compensated for temperature, it is to be expected that the frequency will change with temperature, but a test has not been made to measure the temperature-frequency stability. To insure frequency stability and calibration permanence, particularly at the higher frequencies, it is necessary to make all portions of the oscillator as rigid as possible and to shield the panel. Paper condensers have been found to be more satisfactory than mica condensers.

Alternative Amplitude Control Circuits

The distortion produced by the iron in the amplitude control transformer when a coil of high inductance is used in the tuned circuit can be prevented by taking the amplitude control voltage from the plate circuit of the first amplifier. Unfortunately this involves other difficulties. If capacitive coupling is used, the fluctuations of plate supply voltage of the amplifier stage produce corresponding fluctuations of oscillator grid voltage, which cause the amplitude to vary. If transformer coupling is used, the output of the first stage of the amplifier falls appreciably at low frequencies. The amplitude control compensates for this by increasing the amplitude of oscillation, which also raises the harmonic content. The circuit of Fig. 9, using a coil of not more than three henrys inductance, has proved to be the most satisfactory.

Conclusion

The low harmonic content and high stability of this type of oscillator have made it a valuable addition to the electron tube laboratory.

Appendix

Derivation of Equation for Negative Resistance

In Fig. 10, k represents the fraction of R_b from which voltage is applied to the grid of the opposite tube. Assuming that the two tubes and their supply voltages are in all respects the same, the action is as follows: When the voltage ΔE is zero the points A and B are at the same potential, the two plate currents are equal, and no current flows through the external branch of the circuit. When ΔE is introduced, a current ΔI flows in the external branch. It is composed of two components, ΔI_r, flowing through the plate resistors R_b, and $\Delta I_{p2} = -\Delta I_{p1}$, circulating through the tubes. The resistor current is

$$\Delta I_r = \Delta E/2R_b. \tag{6}$$

ΔE raises the voltage of the plate of tube 2 by $\Delta E/2$ volts and lowers the voltage of the grid of tube 2 by $k\Delta E/2$ volts; the changes of plate and

grid voltage of tube 1 are equal to those of tube 2, but opposite in sign. Assuming that ΔE is sufficiently small so that the plate resistance and transconductance, r_p and g_m are constant, the changes in plate currents are

$$\Delta I_{p2} = \Delta E_p/r_p + g_m\Delta E_g = -\Delta I_{p1} \tag{7}$$

$$= \tfrac{1}{2}\Delta E(1/r_p - kg_m). \tag{8}$$

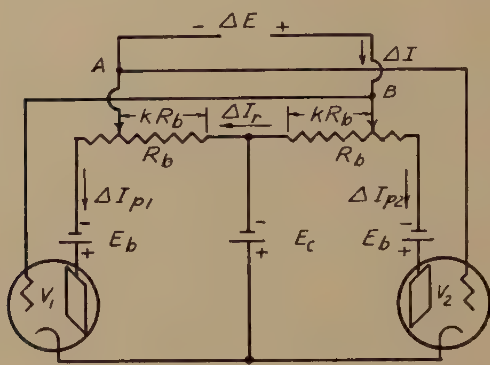


Fig. 10

The total change in I is

$$\Delta I = \Delta I_r + \Delta I_{p2} = \tfrac{1}{2}\Delta E[1/R_b + (1-\mu k)/r_p] \tag{9}$$

$$= \tfrac{1}{2}\Delta E\,\frac{r_p + R_b(1-\mu k)}{r_pR_b} \tag{10}$$

$$= \Delta E/\Delta I = \frac{2r_pR_b}{r_p + R_b(1-\mu k)}. \tag{11}$$

If $k=1$, (11) reduces to (5).

It can be seen from this analysis that the negative resistance of the circuit results from the amplifying property of the tubes. If μ is sufficiently large, then the change in plate current exceeds the change in current through the plate resistors, and is opposite in direction, causing a net external current opposite in direction to the applied voltage.

Proceedings of the Institute of Radio Engineers
Volume 25, Number 11 *November, 1937*

AN ANALYSIS OF ADMITTANCE NEUTRALIZATION BY MEANS OF NEGATIVE TRANSCONDUCTANCE TUBES*

By

E. W. Herold

(RCA Radiotron Division, RCA Manufacturing Company, Inc., Harrison, New Jersey)

Summary—*The reduction of capacitance by feedback from the anode of a negative transconductance tube has frequently been considered an attractive method for compensating for the drop in impedance at high frequencies of a circuit consisting of R and C in parallel. Experimentally, however, the results of previous investigators have never been exceptional. This paper presents a theoretical analysis of the problem on the assumption that an ideal tube is used for admittance neutralization only. The results indicate that approximately flat response of the circuit up to cutoff is obtained when*

$$\frac{C_m}{C_g} = \frac{1}{\sqrt{-2\dfrac{R_p}{R_g}(1+g_mR_p)} - (1+g_mR_p)}$$

where C_m is the feed-back capacitance, C_g is the original shunt capacitance, and R_g is the effective resistance (see Fig. 1(c)). An interdependence between cutoff frequency, C_g, R_g, and the closeness to oscillation is found for a given control-electrode-to-anode transconductance. The effective resistance (and hence the gain of a preceding tube) in the ideal case can be raised several fold over that of the conventional inductance compensated circuit for the same cutoff and shunt capacitance; the greater the increase and consequent improvement, the smaller the safety factor (defined as the fractional increase in transconductance needed to produce oscillation). For the inductance compensated case $\omega_0C_gR_g=1$, while for the tube compensated circuit, the relation is approximately

$$\omega_0C_gR_g = 1.5 + \frac{1}{S}$$

where ω_0 is the cutoff angular frequency and S the safety factor. However, S and R_g are not independent, being related by the approximation

$$g_mR_g = -2\,\frac{S+1}{S^2}\,.$$

A minimum S of about 0.2 would seem reasonable: such a value gives an improvement over inductance compensation of six and one-half times for the ideal tube considered. With a practical tube and circuit, the improvement will probably be much less, possibly one half.

Under ideal conditions, the relations found for the single tube case, apply also to the case of two, cascaded, conventional tubes with feedback from the anode of the second tube to the grid of the first tube.

* Decimal classification: R139. Original manuscript received by the Institute, February 16, 1937. Presented before New York meeting, October 6, 1937.

I. Introduction

In present communication systems, including those suitable for television, a resistance coupled amplifier system is often used to obtain a substantially flat frequency characteristic from a low frequency to a comparatively high one. At the higher frequencies the unavoidable shunt capacitances cause a reduction in the impedance of the coupling elements and so reduce the amplification. There is an ever-present need for some means of compensating for the undesirable capacitances in order to extend the frequency range over which the coupling impedances remain constant in value. The conventional method for compensation is the addition of a small inductance in series with the coupling resistance. Using this method, however, it is found that the resistance which can be used is limited to a definite value which decreases as the cutoff frequency increases. Other methods which appeared to be less limited were frequently proposed but seldom proved to be as satisfactory.

A promising alternative to inductance compensation seemed to lie in the use of feedback from anode to grid of a tube having negative transconductance.[1] Although many studies of feed-back effects in conventional tubes have been made, comparatively few investigators have studied the case of the more unusual tube in which the anode current decreases with an increase in control-electrode voltage; i.e., a change in the positive direction. In the latter type of tube, slight capacitive feedback from anode to grid causes a reduction in input capacitance with a resistive plate load; with a capacitive plate load, a decrease in input conductance is found; with an inductive load, an increase in input conductance will result. These effects are directly opposite to the corresponding ones in conventional tubes.

The first mention of the use of the negative transconductance tube of the retarding field type is in Schottky's classic paper,[2] in which he mentions the usefulness of the negative slope in amplification. In this country, van der Bijl[3] described circuits for a two-grid tube which made use of the phenomenon to obtain a push-pull output from the inner grid and the plate. In such circuits, the feedback to the second (control) grid through the internal capacitance from the inner grid tends to neutralize any change in input admittance caused by capacitive feedback from the plate. Possibly the first attempt to make use

[1] For a discussion of the negative transconductance tube, see E. W. Herold, "Negative resistance and devices for obtaining it," Proc. I.R.E., vol. 23, pp. 1201–1223; October, (1935).
[2] W. Schottky, "High vacuum amplifiers," *Archiv für Elektrotech.*, vol. 8, p. 299; (1919).
[3] H. J. van der Bijl, U. S. Patent No. 1479779, filed 1920.

of this feature was published by Schrader.[4] The latter made use of a two-grid tube with a capacitive load in the inner grid and in the anode. A later study of the effect by Muller[5] showed that the reduction of the input capacitance to the second grid, with a resistive load in the inner grid, could be carried slightly further than the point which neutralized the feedback from the plate. However, instability resulted when too great a reduction of capacitance was attempted. As pointed out by von Ardenne and Stoff,[6] the increase of capacitance which resulted from the use of a two-grid tube instead of a triode might more than make up the slight advantage obtained by overneutralization. Ingenious modifications of the principle have been proposed by others, notably V. D. Landon, P. T. Farnsworth, and W. J. O'Brien.

The writer became interested in the problem of admittance neutralization for wide-band resistance coupling after the development of an improved negative transconductance tube. The new tube appeared to approach the ideal more closely than others previously used. It was felt that this might permit the use of neutralization under conditions which could not be realized in tubes of earlier design. A theoretical investigation of the possibilities was made; this paper presents the more important results. The analysis has been restricted to the case in which the negative transconductance tube is used only for admittance neutralization and not for amplification. It is assumed, therefore, that either a separate tube or a complementary section of the same tube is used as the amplifier succeeding the circuit whose admittance is to be reduced.

II. The Problem

The elementary theory which led investigators to hope for remarkable results with tube compensation is as follows: In a negative transconductance tube (shown by the modified triode symbol in Fig. 1 (a)), when the grid voltage becomes more positive the anode current decreases, and the anode voltage (when a resistive load is used) rises correspondingly. Hence, when an alternating voltage is applied to the grid, the resulting anode voltage will be exactly in phase with it (the cathode is used as a reference point) provided the effect of reactive circuit components is neglected. If the magnitude of the two voltages is equal, any impedance connected from grid to anode would have no

[4] E. Schrader, "On capacitances in electron tubes," *Zeit. für Hochfrequenz.*, vol. 24, p. 35, (1924).

[5] L. Muller, "On the compensation of anode feedback," *Archiv für Elektrotech.*, vol. 16, p. 251; (1926).

[6] M. von Ardenne and W. Stoff, "The compensation of harmful capacitances and their feed-back effects in tubes," *Zeit. für Hochfrequenz.*, vol. 31, pp. 122 and 152, (1928).

alternating voltage across it and no alternating current through it. If the magnitude of the anode voltage is higher, a current, I_m, will flow through the grid-anode impedance *into* the grid in the direction shown by the arrow in Fig. 1 (a). If the magnitude of the alternating anode voltage is of the correct value, this current may be made exactly equal to the current, I_e, which would normally flow into the grid as a result of the applied grid voltage, with feedback absent. The current I_m then completely replaces the current I_e in maintaining the voltage across

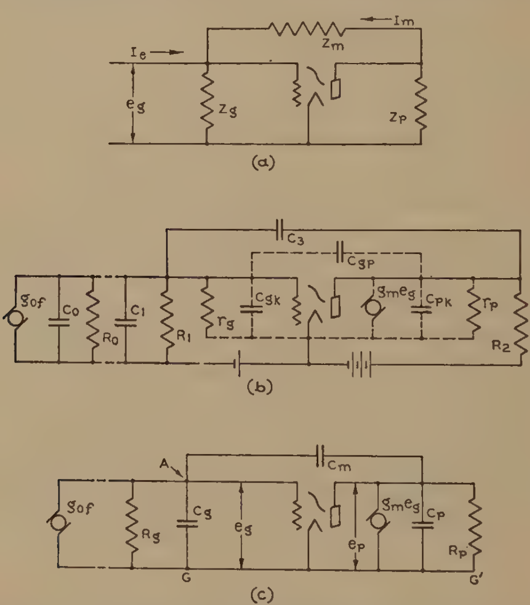


Fig. 1—Equivalent circuits of negative transconductance tube used to neutralize input admittance by feedback through a capacitance.

Z_g, and I_e becomes zero. Under these conditions, the input impedance is infinite. In order to maintain the correct phase relations, provided Z_p is purely resistive, Z_m must have the same phase angle as Z_g.

The simple picture given above is misleading since oscillation sets in when complete neutralization is attempted. To prevent this, it may be reasoned that since it is the effective input capacitance which causes the impedance to drop with frequency, this only need be neutralized. Z_m is then made a pure capacitance which will be called C_m. The resistive part of the input circuit, the reasoning goes, will stabilize the circuit, and the capacitance only will be reduced to zero. Even this situation cannot be realized because of a second complication which enters the picture. The effective anode load consists not only of Z_p

but also of Z_m. Since Z_m is capacitive, even if no capacitance is present directly across Z_p, the plate voltage at high frequencies lags the applied grid voltage somewhat. The result of this lagging voltage is to send a current through Z_m (i.e., C_m) which tends to neutralize the current through the *resistive* branch of Z_g. At some particular high frequency, both the input capacitance and the conductance may be reduced to zero and the system oscillates at or near this frequency. With C_m adjusted to a smaller value than that necessary to assure complete capacitance neutralization, and hence, oscillation, the input impedance will remain finite at all frequencies and self-oscillation cannot occur, provided the circuit elements and the tube characteristics remain constant.

The problem for which a solution is attempted here, is to find the nature of the input circuit impedance with different values of C_m and to determine the usefulness and stability of the feed-back effect in obtaining a flat frequency characteristic. In treating the latter, comparison with the conventional inductance compensated resistive network is made.[7] A signal is supplied by the constant current generator, g_{0f}, with shunting resistance, R_0, and capacitance, C_0.

The basic circuit to be investigated is shown in Fig. 1(b). The negative transconductance tube is represented in its essentials only, by the modified triode symbol.[1] A resistance, R_1, and a capacitance, C_1, comprise the external control-electrode circuit, and a resistance, R_2, the external anode circuit. The contributions of the tube to the circuit including the effect produced by the control-electrode-to-anode transconductance, g_m, are shown in dotted lines. It is assumed that the inverse, or anode-to-control-element, transconductance is negligible; i.e., the control electrode draws no current which is affected by the anode potential. For purposes of analysis it is convenient to lump together parallel resistances and capacitances. This has been done in Fig. 1(c). In the simplified circuit r_p and R_2 are lumped together as R_p; R_0, r_g, and R_1 are lumped as R_g and the various capacitances are lumped in the three, C_g, C_p, and C_m, as indicated. The operation of the circuit is completely determined by calculation of the total admittance from points A to G of Fig. 1(c); this admittance will be designated by Y_{AG}.

III. The Analysis

The fundamental relationship of the constants of Fig. 1(c) which will result in oscillation, must be determined before going further. This relationship is derived by application of Kirchhoff's laws to the

[7] The conventional inductance compensation consists of the addition of a small inductance in series with the resistive branch of the $C-R$ combination.

circuit[8] and by postulating no external applied electromotive force; i.e., $f=0$. The total current entering the lower wire $G-G'$ of Fig. 1(c) is

$$e_g\left(\frac{1}{R_g}+j\omega C_g+g_m\right)-e_p\left(\frac{1}{R_p}+j\omega C_p\right)=0$$

where e_g is the difference in potential between grid and cathode, and e_p is the difference in potential between cathode and anode. The current leaving the point A in the figure is

$$e_g\left(\frac{1}{R_g}+j\omega C_g+j\omega C_m\right)+e_p(j\omega C_m)=0.$$

These equations give as the oscillatory solution

$$\begin{vmatrix} \left(\frac{1}{R_g}+j\omega C_g+g_m\right) & -\left(\frac{1}{R_p}+j\omega C_p\right) \\ \left(\frac{1}{R_g}+j\omega C_g+j\omega C_m\right) & j\omega C_m \end{vmatrix}=0.$$

Equating the imaginary and real parts of the determinant to zero gives the conditions

$$\omega_{\mathrm{osc}}=\sqrt{\frac{1}{R_gR_p(C_gC_p+C_gC_m+C_pC_m)}} \tag{1}$$

and

$$C_{m_{\mathrm{osc}}}=-\frac{C_gR_g+C_pR_p}{R_g+R_p+g_mR_gR_p} \tag{2}$$

which may be solved for g_m to give

$$g_{m_{\mathrm{osc}}}=-\left[\frac{\frac{C_g}{C_m}+1}{R_p}+\frac{\frac{C_p}{C_m}+1}{R_g}\right]. \tag{3}$$

The subscript, osc, on ω, C_m, and g_m refers to the values which just cause oscillation. The oscillations which occur are sinusoidal at small amplitudes. Equations (1), (2), and (3) are closely checked by experiment.

The admittance from points A to G of Fig. 1(c) is the usual expression

$$Y_{AG}=\frac{1}{Z_g}+\frac{1+g_mZ_p}{Z_m+Z_p} \tag{4}$$

where $Z_m=1/j\omega C_m$ and Z_g and Z_p include the C_g-R_g circuit and the C_p-R_p circuits, respectively.

[8] The author derived this relationship in the paper referred to in footnote 1.

By calculation of the modulus of Y_{AG} for various values of C_m and ω, with the other circuit constants given, the problem is solved formally. It will be found that a value of C_m may be chosen which gives an admittance approximately independent of frequency up to a certain cutoff point. When this value is used in (3), the transconductance to cause oscillation may be found and used to determine how close the circuit is to the point of oscillation. The percentage difference between $g_{m_{osc}}$ (the value necessary for oscillation) and g_m (the normal value) gives a safety factor for the circuit. The improvement in gain over an inductance compensated case can also be estimated. A typical case has been calculated and the results are shown in Figs. 2 and 3.

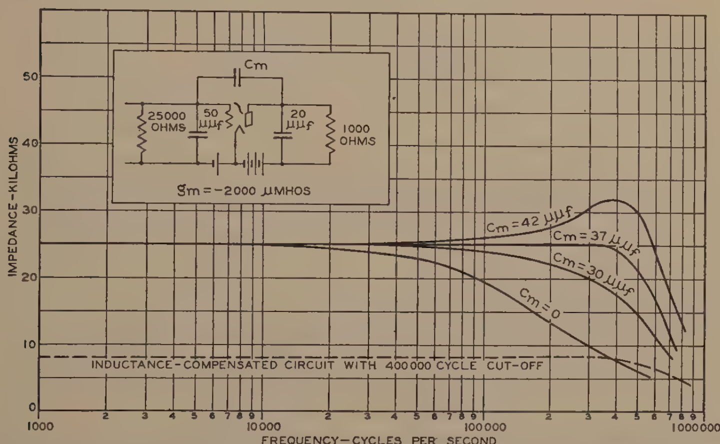


Fig. 2—Modulus of input impedance plotted against frequency for a negative transconductance tube with various feed-back capacitances.

The equations which result from an attempt to carry out this procedure in a completely general case are so unwieldy as to make it virtually impossible to obtain useful information for comparison of the tube method with the inductance compensated method. It has been found that by assuming an ideal tube and circuit, in which $C_p = 0$, the merits of the circuit can be evaluated.

When C_p is zero, the admittance becomes

$$Y_{AG} = \frac{1}{R_g} + j\omega C_g + \frac{1 + g_m R_p}{R_p + \dfrac{1}{j\omega C_m}}$$

$$= \frac{1}{R_g} + \frac{\omega^2 C_m{}^2 R_p (1 + g_m R_p)}{1 + \omega^2 C_m{}^2 R_p{}^2} + j\omega C_g \left(1 + \frac{C_m}{C_g}\,\frac{1 + g_m R_p}{1 + \omega^2 C_m{}^2 R_p{}^2}\right). \quad (5)$$

This equation shows clearly that both conductive and susceptive parts of the input are reduced by feedback even though $C_p = 0$. Flat response may be considered to be given by the value of C_m which leads to a slowly rising value of the admittance as the frequency is increased, with no minima. To find this condition, the modulus of the admittance, $|Y_{AG}|$, may be differentiated with respect to ω and the frequencies found at which maxima or minima exist by equating the derivative to zero. When ω is set equal to zero in the result (the $\omega = 0$ root having first been eliminated), all maxima or minima occur at zero frequency. The condition given is then that for which the minimum in the admit-

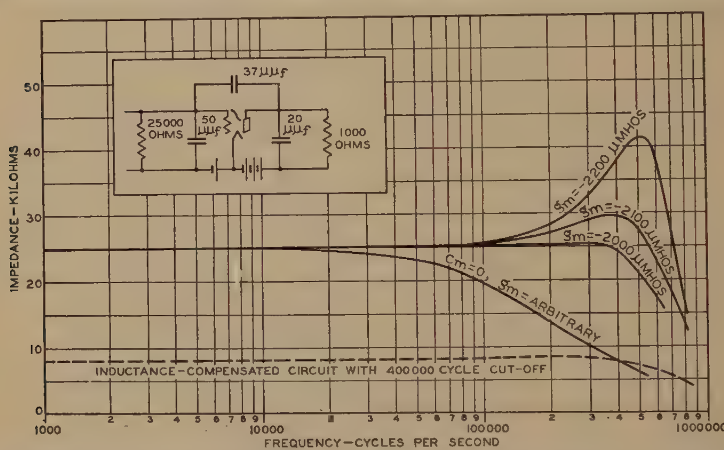


Fig. 3—Modulus of input impedance plotted against frequency for a negative transconductance tube with fixed feed-back capacitances and various tube transconductances.

tance versus frequency curve just disappears. The value of C_m given will, therefore, be that for flat response.

To perform the differentiation, note, that if

$$Y_{AG} = a + jb$$

then,

$$|Y_{AG}| = \sqrt{a^2 + b^2}$$

and

$$\frac{d|Y_{AG}|}{d\omega} = \frac{a\dfrac{da}{d\omega} + b\dfrac{db}{d\omega}}{|Y_{AG}|}.$$

The condition for this equation to be zero is

$$a\frac{da}{d\omega} + b\frac{db}{d\omega} = 0.$$

Using this method on (5) gives

$$\left(\frac{1}{R_g}+\frac{r^2x^2R_pk}{1+r^2x^2R_p^2}\right)\left[\frac{(1+r^2x^2R_p^2)2r^2kR_p-2r^4x^2kR_p^3}{(1+r^2x^2R_p^2)^2}\right]x$$
$$+\left(1+\frac{rk}{1+r^2x^2R_p^2}\right)\left[\frac{-2r^3kx^2R_p^2}{(1+r^2x^2R_p^2)^2}+\left(1+\frac{rk}{1+r^2x^2R_p^2}\right)\right]x=0 \tag{6}$$

where,

$$r=\frac{C_m}{C_g}, \qquad k=(1+g_mR_p), \qquad x=\omega C_g.$$

Taking out the solution $\omega=0$ (corresponding to $x=0$), and setting $\omega=0$ in the other solution, gives

$$\frac{2r^2kR_p}{R_g}+(1+rk)^2=0.$$

Hence,

$$r=\frac{1}{-k+\sqrt{-2kR_p/R_g}}$$

and[9]

$$\frac{C_m}{C_g}=\frac{1}{\sqrt{-2\frac{R_p}{R_g}(1+g_mR_p)}-(1+g_mR_p)}. \tag{7}$$

If C_m is set to this value, it is now of interest to calculate the transconductance which will cause oscillation. Use of (3) and (7) gives (on the continued assumption of an ideal tube with $C_p=0$)

$$g_{m_{osc}}=g_m-\frac{1}{R_g}-\sqrt{\frac{-2(1+g_mR_p)}{R_gR_p}}. \tag{8}$$

Since the safety factor, S, is defined as

$$S=\frac{g_{m_{osc}}}{g_m}-1$$

we have

$$S=-\frac{1}{g_mR_g}\left[1+\sqrt{-2\frac{R_g}{R_p}(1+g_mR_p)}\right]. \tag{9}$$

In using these relations to plot the admittance, Y_{AG}, it is found that the cutoff frequency is approximately the same as that frequency at

[9] The physically valid solution is used.

which the susceptance of (5) becomes equal to $1/R_g$. This is, of course, reasonable since the susceptance at this point rises rapidly as the frequency increases with a consequent reduction in impedance at higher frequency. Since $\omega^2 C_m{}^2 R_p{}^2$ is always small at the cutoff point, it may be neglected in comparison with unity without serious error. If ω at cutoff is called ω_0, the relation is

$$\omega_0 C_g\left[1 + \frac{C_m}{C_g}(1 + g_m R_p)\right] = \frac{1}{R_g}$$

and

$$\omega_0 C_g R_g = \frac{1}{1 + \dfrac{C_m}{C_g}(1 + g_m R_p)}. \tag{10a}$$

From[9] (7)

$$\omega_0 C_g R_g = 1 + \sqrt{\frac{-R_g}{2R_p}(1 + g_m R_p)}. \tag{10b}$$

Equations (7), (9), and (10a) or (10b) are the complete answer to the problem since by elimination between them, R_g and R_p can be found in terms of a given g_m, C_g, safety factor, and cutoff frequency. However, experience indicates that the two equations which result are misleading because they sometimes indicate a solution which is physically either impossible or impracticable. For this reason it is believed best, for design purposes, to use (7), (9), and (10a) or (10b) in the form given.

Equations (10a) and (10b) are in such form that the right-hand side represents directly the improvement factor over the inductance compensated case[10] in which $\omega_0 C_g R_g = 1$. In Fig. 4, the improvement factor is plotted against R_p for $g_m = -2\times 10^{-3}$ mhos and for two values of safety factor, namely, $S = 0.2$ and $S = 0.3$. It is seen that substantial improvement over the conventional method is realized only for the higher values of R_p. The dotted curves show $\omega_0 C_g$ for the same two values of S. The variation of $\omega_0 C_g$ with R_p, when the latter is greater than 1000 ohms, is seen to be small. The use of a reasonably high value for R_p is, therefore, doubly warranted: it entails no great sacrifice in cutoff and it gives greater improvement over the inductance compensated circuit.

By substituting (7), (9), and (10) in (5), it can be shown that under the flat response condition

[10] See G. V. Braude, *Jour. Tech. Phys.* (U.S.S.R.), vol. 4, pp. 1714–1739 and pp. 1818–1828, (1936), and O. Lurje, *Tech. Phys.* (U.S.S.R.), vol. 3, pp. 229–248, (1936). The conditions for flat response in the inductance compensated case are $R^2 = 2.414L/C$ and cutoff is approximately at $\omega_0 CR = 1$.

$$Y_{AG} = \frac{1}{R_g}\left[1 - \frac{1}{2}\frac{\omega^2}{\omega_0{}^2}\frac{1}{1 + \dfrac{\omega^2}{\omega_0{}^2}\dfrac{1}{4(\omega_0 C_g R_g - 1)^2}} + j\frac{\omega}{\omega_0}\left(\frac{1 + \dfrac{\omega^2}{\omega_0{}^2}\dfrac{\omega_0 C_g R_g}{4(\omega_0 C_g R_g - 1)^2}}{1 + \dfrac{\omega^2}{\omega_0{}^2}\dfrac{1}{4(\omega_0 C_g R_g - 1)^2}}\right)\right]. \tag{11a}$$

If S is not too great and the higher values of R_p are used so that $\omega_0 C_g R_g$ is appreciably greater than unity and if ω/ω_0 is restricted to the useful range, (11a) simplifies to

$$Y_{AG} = \frac{1}{R_g}\left[1 - \frac{1}{2}\frac{\omega^2}{\omega_0{}^2} + j\frac{\omega}{\omega_0}\right]. \tag{11b}$$

This equation is useful to plot the frequency response as in Fig. 5 where the conductance, the susceptance, and the modulus of the impedance are plotted against ω/ω_0.

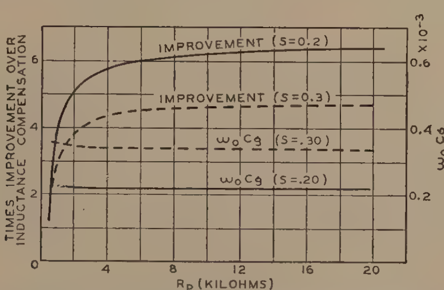


Fig. 4—Relationship between improvement over inductance compensated case, susceptance at cutoff, and anode-load resistance for feedback giving flat response. Two values of safety factor are used and $g_m = -2\times10^{-3}$ mhos.

Since it is clearly indicated that R_p should be high, it is of some importance to examine the design equations when $g_m R_p \gg 1$. Making this assumption in (9) and (10b) gives for the upper limit of improvement over inductance compensation

$$\omega_0 C_g R_g = 1 + \sqrt{\frac{S + 1 + \sqrt{2S+1}}{2S^2}} \tag{12a}$$

$$\approx 1.5 + \frac{1}{S} \text{ (less than 3\% error for } S < 0.6). \tag{12b}$$

This result is, therefore, the best that the tube neutralized circuit is capable of attaining. It must be remembered, however, that (12) is

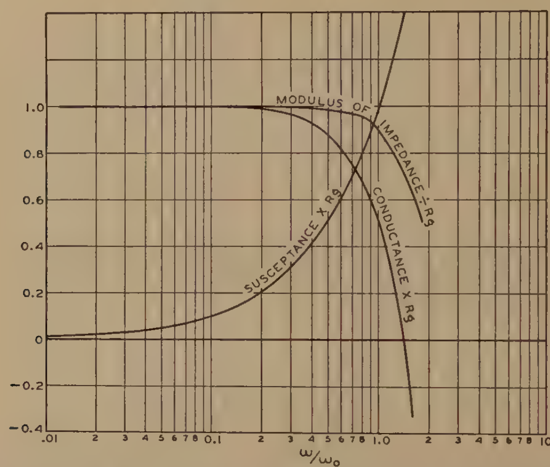


Fig. 5—Curves showing variation of modulus of input impedance and the conductive and susceptive components with frequency. High anode-load resistance is assumed. The curves are plotted with dimensionless variables so as to preserve generality.

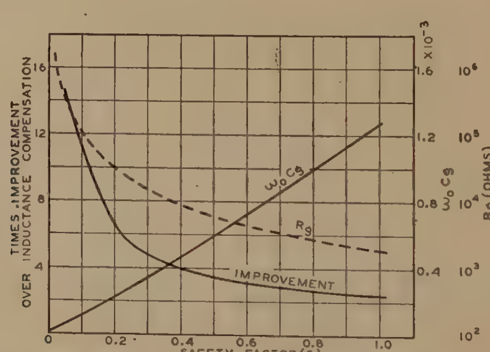


Fig. 6—Curves showing important design factors vs. safety factor, assuming $g_mR_p \gg 1$ and with $g_m = -2 \times 10^{-3}$ mhos.

not the complete picture: it is not possible to choose S and R_g independently if g_m is also given. The relation between S, g_m, and R_g is

$$g_mR_g = -\frac{(S+1)+\sqrt{2S+1}}{S^2} \tag{13a}$$

$$\approx -2\frac{S+1}{S^2} \text{ (less than 4\% error for } S < 0.6). \tag{13b}$$

Fig. 6 has been plotted from (12a) and (13a), again on the assumption that $g_m = -2 \times 10^{-3}$ mhos. In the figure it is seen that $\omega_0 C_g$ varies almost linearly with S. R_g (which is shown on a logarithmic scale) falls rapidly as S and $\omega_0 C_g$ are increased. The dotted curve shows the ratio of grid impedance to that possible with inductance compensation for the same cutoff as given by $\omega_0 C_g R_g$.

With a practical tube, a lower limit for the safety factor might be placed at 0.2 to 0.3. For the former value, the maximum improvement of the tube over the inductance method is 6.5; for the latter value it is nearly 5. In all actual cases, the presence of C_p alters the situation to the detriment of the improvement over inductance compensation. The general effect is a lowering of the R_g curve as well as the improvement factor curve of Fig. 4. As an example, in the case of Fig. 2, the improvement in impedance over the inductance compensated circuit is about three times. Equation (10b), however, would lead one to believe the improvement to be about 4.5 times.

Another consequence of assuming an ideal tube in deriving the equations is that apparently no increase in advantage over inductance compensation would be found by increasing g_m; in order to realize the same safety factor, R_g must be decreased in roughly the same proportion. $\omega_0 C_g$ is increased but the improvement over the inductance compensated case is the same. In all practical cases, a high g_m is necessary in order that $g_m R_p$ may be made high without increasing R_p to such a point that the shunting effect of C_p becomes appreciable.

In passing, it may be mentioned that the effect of C_p on the plate load may be compensated in part by inductance in series with the plate resistor. This should enable a practical amplifier to approach the ideal case above more closely.

IV. Word Picture and Summary

A word picture of the results obtained above as they apply to a tube neutralized admittance is of some value in clarifying the analysis. It is first necessary to recall the preliminary picture of the operation given in stating the problem. With a negative transconductance tube having a resistive load in the anode, the alternating anode potential has the same phase with respect to the cathode as the control-electrode potential producing it. When a small capacitance, C_m, is connected from anode to control electrode, some of the alternating anode current flows into the input circuit and leads the input voltage in phase. By the elementary theory, this leading current reduces the effective capacitance in the input from its original value to some lower value. A more rigorous analysis shows that the feed-back capacitance, C_m, is part of

the anode load and causes the anode voltage to lag the control-electrode voltage slightly, the lag being roughly proportional to frequency. The lagging component of voltage sends a current, opposite in phase to the input voltage, through C_m to the input circuit. The effective input conductance of the tube is then negative and subtracts from the original conductance of the input circuit.

As a result, the total admittance consists of two parts, a conductance and a susceptance. The conductance consists of the original conductance less the conductance due to feedback. Since the capacitance is also the original less that due to feedback, the susceptance is reduced to some fraction of the original susceptance before feedback. There is, however, an important difference in the reduction of the two components; the susceptance is reduced by a fraction which is roughly independent of frequency (it depends on the transconductance and the feed-back capacitance), whereas the negative conductance due to feedback depends on the lagging component of the anode voltage divided by the feed-back impedance, and hence increases with the square of the frequency. As the frequency is increased, the total conductance is reduced until it passes rapidly through zero and becomes negative (see Fig. 5). While the susceptance could also be reduced to zero by sufficient feedback, only by retaining a positive susceptance which increases with frequency, is oscillation avoided. The susceptance which must be retained is higher, for higher values of the safety factor (defined as the fractional increase in transconductance necessary to cause oscillation), as would be expected. Cutoff is defined by the frequency at which the susceptance becomes equal to the zero-frequency conductance (which is also the original conductance before feedback). This cutoff point depends, therefore, on the safety factor.

In applying the results to an actual design, let it be assumed that the gain to the anode ($g_m R_p$) can be made reasonably great, since this is desirable. The safety factor chosen determines the product of the grid resistance which can be used and the transconductance (equation (9)). The latter product, in turn, determines the grid resistance, since the transconductance is usually known. The grid resistance then fixes the susceptance of the remaining capacitive component in the input at cutoff (equation (10a)); since the actual value of the remaining capacitance is also determined by the grid resistance and the transconductance (equations (5) and (7)), the cutoff frequency is definitely fixed (equation (10b)). The frequency response (equation (11a)) may then be plotted, if desired.

V. Note on Two-Tube Circuit

Among other arrangements, one frequently suggested for input admittance neutralization is a two-tube, cascaded, resistance coupled amplifier with a feed-back capacitance from the second tube anode to the first tube grid. The writer believes that the performance of a two-stage amplifier of this kind cannot be easily treated by theoretical methods for superaudible frequencies because of the unavoidable and complex degenerative and regenerative effects in practical amplifiers. However, assuming that screen-grid tubes are used (so that the Miller effect can be neglected), it is possible to obtain, by a rough analysis based on analogous reasoning to that used in this paper, some information as to performance. Preliminary investigation indicates that the two-tube circuit, properly designed, may approach closely in performance the ideal single tube case treated above.

VI. Conclusion

By the theoretical analysis, it would appear that input admittance neutralization by means of feedback from the anode of a negative transconductance tube has strictly limited possibilities. The proximity of the circuit to oscillation depends on the cutoff frequency and the value of the input admittance in the pass band. The improvement of such a circuit over the conventional inductance compensated one may be measured by the ratio of their impedances for a given cutoff frequency and is greater when the operation is closer to the oscillation point in the tube case. For reasonable safety factors, an ideal tube would provide improvements of five to six times. With practical tubes, the improvement is much less, possibly one half of that for an ideal tube.

Proceedings of the Institute of Radio Engineers
Volume 25, Number 11 *November, 1937*

ON THE IONIZATION OF THE F_2 REGION*

By
W. M. Goodall
(Bell Telephone Laboratories, Inc., Deal, New Jersey)

Summary—*In this paper the available data on F_2 region ionization for Peru, Australia, and this country are analyzed in a way that permits the separation of effects due to variations in solar ionizing force from effects due to seasonal and annual changes. It is shown that for constant solar activity the expected curves of critical frequency for Australia and this country appear to indicate both seasonal and annual tendencies. It is suggested as a possibility that the apparent "annual" effect may in fact be due to meteorological conditions which cannot be eliminated without data from more locations.*

FOR some time now it has been known that the critical frequencies of the F_2 region do not vary in the same manner as those for the lower regions.[1] This is true, not only for the diurnal, but also for the yearly variations. When data obtained only at Washington and Deal were available, it was possible to explain the results as a seasonal effect involving "heating" of the F_2 region.[2] Later results obtained in the southern hemisphere by the Carnegie Institution indicate, as was pointed out by Berkner, Wells, and Seaton,[3] that the effect appears to be annual rather than seasonal in nature.

It seems to me, however, that it may be too early to conclude that an annual effect is predominant. It is interesting that measurements were started at Watheroo, Australia, during a period of increasing solar activity. If solar activity is an important factor in the F_2 ionization,[4] as it appears to be, measurements obtained during a limited time when the activity is changing might not be representative of conditions for constant activity. The height data, on the other hand, suggest that there may be a seasonal effect.

Because of these uncertainties it is desirable to analyze the avail-

* Decimal classification: R113.5. Original manuscript received by the Institute, April 15, 1937. Presented before Eighteenth Annual Meeting, American Geophysical Union, Washington, D. C., April 28, 1937.

[1] Schafer and Goodall, *Nature*, vol. 132, p. 522; September 30, (1933). Kirby, Berkner, and Stewart, "Studies of the ionosphere and their relation to radio transmission," Proc. I.R.E., vol. 22, p. 508; April, (1934).

[2] Theories of this type have been proposed by E. V. Appleton and E. O. Hulburt. See E. V. Appleton and R. Naismith, "Some further measurements of upper atmosphere ionization," *Proc. Roc. Soc.*, ser. A, vol. 150, pp. 685–708; July, (1935), and E. O. Hulburt, "Theory of the ionosphere," *Terr. Mag.*, vol. 40, pp. 193–200; June, (1935).

[3] Transactions of the American Geophysical Union, Seventeenth Assembly, 1936.

[4] Ionization or ionic density is proportional to the square of the critical frequency.

able data in such a way as to eliminate both effects. This may be done by treating the twelve months separately; for each month we assemble the available data and plot it against an appropriate measure of solar

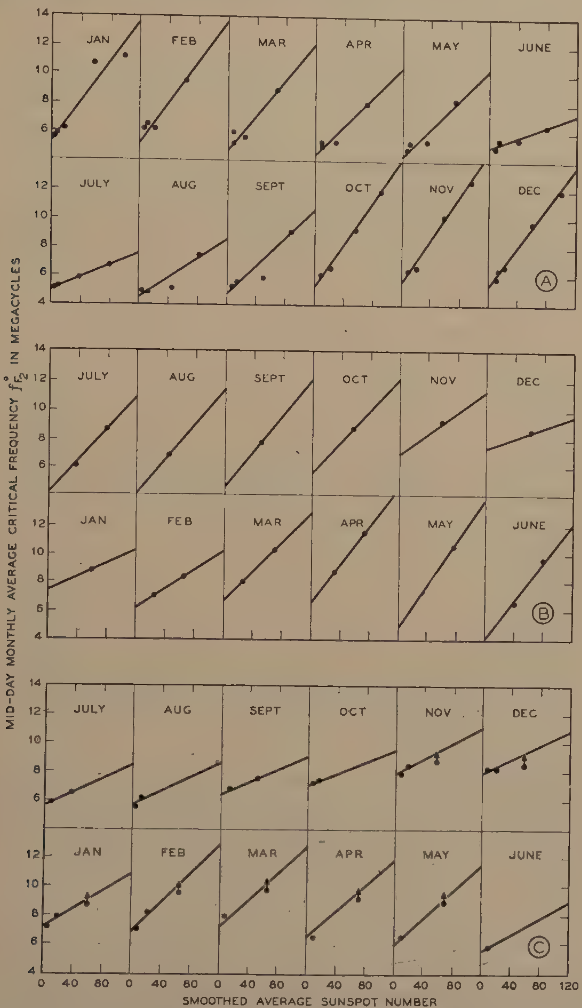


Fig. 1—Monthly correlation of critical frequencies and smoothed sunspot numbers. A, Washington-Deal; B, Watheroo; C, Huancayo.

activity.[5] The measure used is the smoothed or twelve month running average value of the relative sunspot number. The results are shown in

[5] Correlation of *annual* average values of midday F_2 region critical frequencies with sunspot numbers has been shown by E. B. Judson, Proc. I.R.E., vol. 25, p. 43; January, (1937).

Fig. 1. The data for Fig. 1A were obtained at Deal, N. J., for 1933[6] and at Washington, D. C., for 1934 to 1937.[7] The data for Fig. 1B were obtained at Watheroo, Australia, during 1935 and 1936 and for Fig. 1C, Huancayo, Peru, during 1934–1935 and 1936.[8] It is seen that for Washington and Deal the correlation is reasonably good. The data for Australia and Peru are not as complete, but it is seen that they are consistent with a solar effect, for those months where sufficient data are available.

Another way of showing the correlation, once the curves of Fig. 1 have been obtained, is shown in Fig. 2 where the smooth curve is what would have been expected from smoothed sunspot numbers on the basis of the straight line curves of Fig. 1A. The points are observed values.

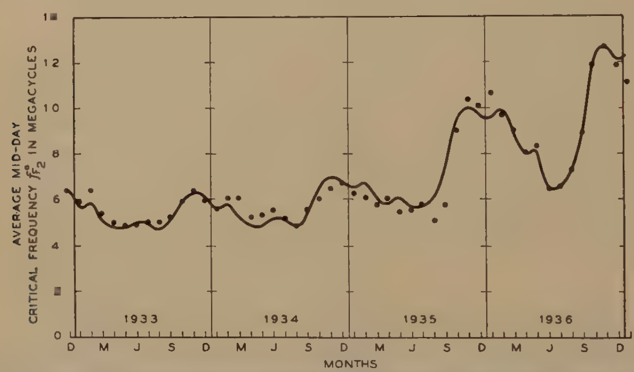


Fig. 2—The curve gives the expected values on the basis of the curves of Fig. 1A. The circles are observed values (Washington-Deal).

We next assume different sunspot numbers as parameters and obtain from Fig. 1 the values of the midday critical frequencies normally to be expected for different months. In these curves (Fig. 3) we have the trend of the F_2 critical frequency to be expected for constant conditions of solar activity. For each location we have three curves for sunspot numbers R of 10, 50, and 100 corresponding to low, moder-

[6] The Deal data have been published in the Proc. I.R.E., vol. 23, p. 674; June, (1935).

[7] The Washington data were obtained by the National Bureau of Standards, See *Terr. Mag.*, vol. 41, pp. 379–388; December, (1936), and current *Science Service* bulletins.

[8] The data for Watheroo and Huancayo were obtained by the Carnegie Institution. See publication number 1481 of the Department of Terrestrial Magnetism, and *Terr. Mag.*, vol. 41, pp. 173–184; June, (1936). For corresponding *seasons* the slopes of the Watheroo curves for months when two points were available are nearly the same as those of the Washington curves. For this reason the slopes of the other Watheroo curves were made the same as those of the corresponding Washington curves.

ate, and pronounced activity. The curves are plotted on a seasonal basis; local summer is at the middle of each and local winter at the

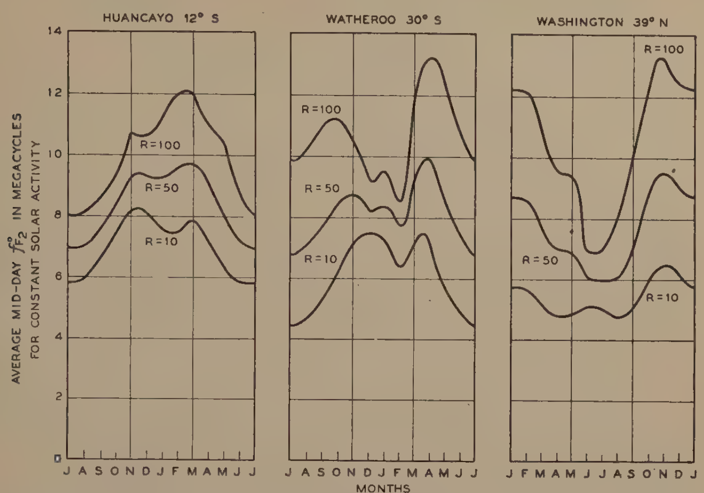


Fig. 3—Expected values of average midday critical frequencies for constant solar activity. R is the smoothed sunspot number.

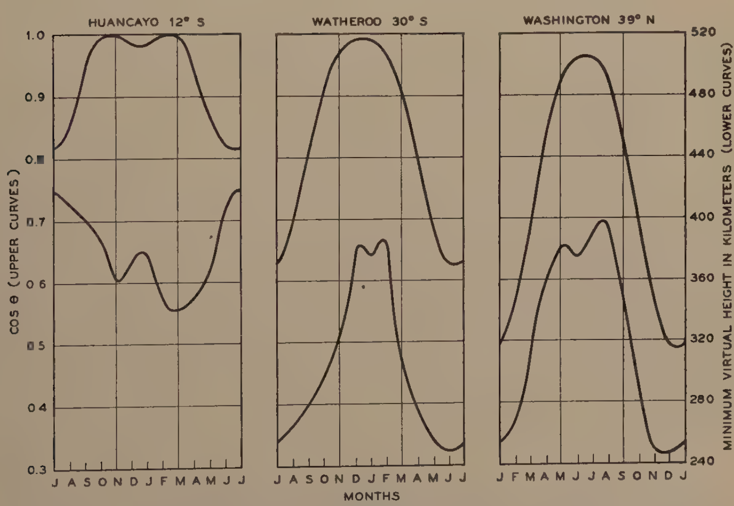


Fig. 4—Curves of the sun's zenith angle at noon (upper curves), and of minimum virtual height (lower curves).

ends. Seasonal tendencies for corresponding northern and southern latitudes would be indicated by curves of similar shape. Annual tendencies would be indicated by similar shapes for the same months, instead of for the same seasons.

In these curves there appear to be tendencies for both types of effect. Thus, for both Washington and Watheroo there is a tendency for a minimum to occur during local summer. On the other hand, for Watheroo there is a somewhat broader tendency for a maximum to occur during the same months as at Washington. The former tendency is seasonal in nature while the latter is annual. Similar comparisons between Huancayo and the other locations can be made, but the significance of these comparisons is not clear cut because this station is at a different latitude from the others and the deviations could be due to latitude effects.

There seems to be some relation between the curves of Fig. 3, the sun's zenith angle and the height of the F_2 layer. In Fig. 4 the upper curves are the values of the cosine of the sun's zenith angle at local noon while the lower curves are the average of the observed values of the minimum virtual height of the F_2 region. The ionizing force, other things being equal, should be roughly proportional to the upper curves. The lower ones must be related to a number of factors, including some of significance in the meteorology of the lower atmosphere. It is natural to assume that higher F_2 regions signify higher temperatures at a given time of day, as at noon. It is likely that for a given ionizing force higher temperatures result in lower ionic densities. If we (perhaps rather uncritically) make these assumptions, a plausible explanation of the critical frequency curves of Fig. 3 can be made on the basis of the curves of Fig. 4; high ionic density being due either to high ionizing force or to low temperature, and low ionic density to low ionizing force or to high temperature. This is all qualitative; it is of interest as a possible explanation of the observed results and as a suggestion for further research. In particular, this is mentioned because it suggests that perhaps we should canvass other possibilities before concluding that an annual effect exists. Is it not possible that this "annual effect" is really due to variations with latitude or general climate which cannot be eliminated because of the small number of places where we have taken data?

Proceedings of the Institute of Radio Engineers

Volume 25, Number 11 *November, 1937*

ELECTROMAGNETIC WAVE FIELDS NEAR THE EARTH'S SURFACE*

By

Charles R. Mingins

(Department of Physics, Cornell University, Ithaca, New York)

Summary—*It is shown that in the case of frequencies where the ground wave is of importance any indirect waves can be investigated by keeping within the "sky-wave furrow" of the directional pattern of a loop receptor. Observations indicate multiple downcoming waves in the daytime, with marked changes occurring toward sunset. Conditions during the day are as a rule steady and sometimes resume their steadiness after the sunset fluctuation period.*

During steady conditions observations were made of the field in the neighborhood of various systems involving boundaries between media of different conductivities, such as horizontal and vertical conductors near the earth's surface and the gorges which characterize the terrain in and about Ithaca. The results are interpreted on the basis of a steady field upon which is superimposed a perturbation field due to a conductor or anticonductor of conductivity $\pm\Delta\sigma$.

Several ways in which the wave plurality may be accounted for are discussed.

IT HAS become rather firmly established as a fact in the transmission of electric waves that the energy from the transmitter reaches the receiver in general partly by a direct path along the ground and partly by an indirect path or paths, being deflected down from higher regions. This multiplicity of routes gives rise to a complicated field pattern at the receiver, because of the mutual superposition of the "ground wave" and one or more "sky waves."

The first task, then, of one who wishes to study downcoming waves, in dealing with any frequency for which the ground wave is important at the given distance from the transmitter, is to eliminate the ground wave. Choice might be made of frequencies high enough to eliminate the ground wave as a factor, but in the so-called "broadcast range" the experimental difficulties are somewhat less. The occurrence of steady conditions, for observations of a "control" nature and for other purposes of the present investigation as discussed below is far more frequent in this range than at higher frequencies. The desirability of a method which uses regular program transmissions, thereby avoiding the necessity for the operation of transmitting apparatus, is obvious.

* Decimal classification: R113. Original manuscript received by the Institute, March 15, 1937. A thesis presented to the Faculty of the Graduate School of Cornell University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. More complete mathematical details, tables, etc., are to be found in the original thesis on file in the Cornell University Library.

Various devices have been resorted to by experimenters in order to balance out the ground wave, involving more or less manipulation, as, for instance, the introduction of a definite amount of ground-wave electromotive force, in opposite phase, from an extraneous receptor.

A partial separation of sky wave from ground wave was attained by Merritt[1] and co-workers by observing the phase differences between that component of the sky-wave magnetic vector along a certain definite line, the intersection of the plane of incidence with the surface of the ground, and a component of the composite effect of ground- and sky-wave magnetic forces. From this separation valuable evidence[2] concerning the nature of the complex interference phenomena connected with fading was obtained.

The present study concerns itself with effects due to the composition of the sky-wave horizontal magnetic component mentioned above with the sky-wave vertical magnetic component; that is, it extends the scope of observation of downcoming radiation to the resultant effect in a plane in which sky wave alone acts.

Specifically, consider the pattern obtained with a receiving coil aerial, or loop or frame, by plotting received ground-wave intensity against angular direction of the normal to the plane of the loop. If the coil is swung about a vertical axis lying in the plane of the coil (an axis in the plane of the coil will hereafter be termed "the axis," as distinguished from the normal to the plane of the loop), a figure-of-eight pattern is obtained, the received impulse being zero when the normal lies in the plane of incidence, and a maximum when the plane of the coil coincides with the plane of incidence. Likewise, if the axis of rotation, fixed in the coil as described, is given any orientation in the plane of incidence while a rotation of the coil about it of 2π radians is made, a similar figure will be obtained, with the zero position always occurring when the normal to the coil lies in the plane of incidence. The complete resulting pattern for the ground wave is, therefore, a solid figure of eight divided squarely in two at zero intensity by the plane of incidence, if a directed normal be used for reference.

If the axis of the loop, therefore, is set so as to be perpendicular to the plane of incidence, no matter how the loop is then rotated no ground wave can be received. Any actuation of the receiver must then be due to sky wave, and we may term the "furrow" around the middle of the solid figure of eight the "sky-wave furrow."

It was earlier believed by some[3] that the maximum signal would

[1] Merritt and Bostwick, *Proc. Nat. Acad. Sci.*, vol. 14, p. 884; November, (1928).

[2] Merritt, McLean, and Bostwick, *Jour. Frank. Inst.*, vol. 211, p. 539; May, (1931).

[3] *Vide*, e.g., Eccles, *Proc. Roy. Soc.* ser. A, vol. 87, p. 96; August, (1912).

be received at a given place on the earth by a linear antenna when it was made to assume the position of the electric force in the downcoming wave front. For the case of a coil aerial it has been stated[4] also that the direction of incidence should be observable directly at a station on the ground as the direction of the normal to the loop (in the plane of incidence) for zero received impulse. These implications that direction finding can be carried out directly on a downcoming wave neglect the presence of the reflector—the surface of the earth—near which the receiver is located, and which gives rise to a reflected wave

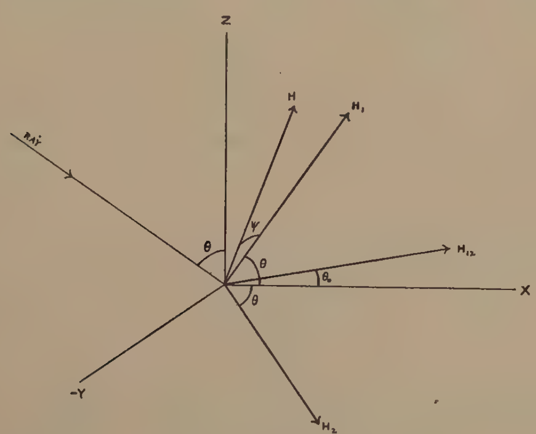


Fig. 1

interfering with the incident one. In regions which have very moist climates, or in which the ground is habitually damp, as appears to be the case in England,[5] the resultant electric force will not appreciably depart from the vertical or the magnetic force from the horizontal.

Let us examine the general situation by reference to Fig. 1. Since we are to use a loop as receptor, it will be simplest to consider the magnetic force in the wave. We shall choose right-handed Cartesian co-ordinate axes, letting z be along the vertical, positive upwards, and the xz plane the plane of incidence. Let θ be the angle of incidence of the downcoming ray; θ will then be also the angle of reflection. If we let H be the instantaneous magnetic vector of the incident ray, assuming for the moment a plane polarized wave, and that H makes an angle ψ with the plane of incidence, then

$$H_1 = H \cos \psi,$$

H_1 being the component of H in the plane of incidence. It is obvious

[4] Austin, Proc. I.R.E., vol. 13, pp. 409–412; August, (1925).

[5] E.g., Appleton and Barnett, *Proc. Roy. Soc.*, ser. A, vol. 109, p. 630; December, (1925).

that H_1 makes an angle θ with the x axis. Then H_2, the magnetic component of the reflected wave in the plane of incidence, as shown also

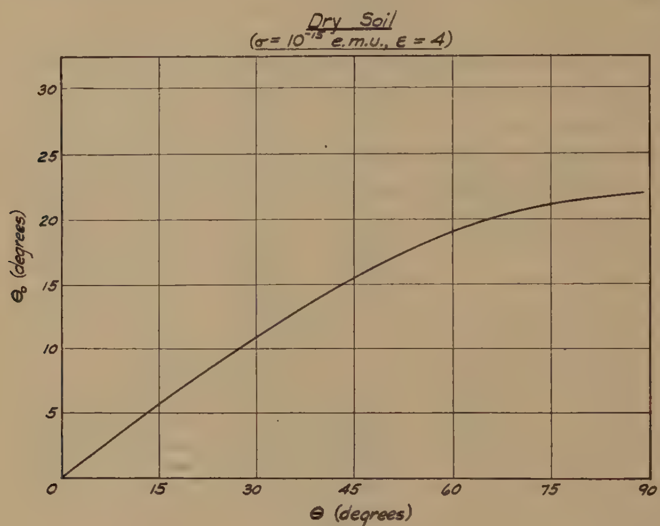


Fig. 2

makes an angle θ with the x axis, and if the reflection is perfect the resultant of H_1 and H_2 is horizontal. This would mean a sky-wave figure of eight in the sky-wave furrow with its null line vertical, and

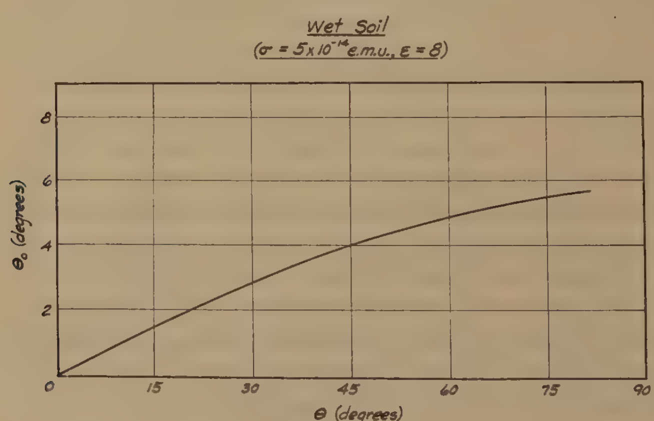


Fig. 3

one would under such circumstances always read the apparent angle of incidence as zero, no matter what the actual direction of approach of the waves.

If the ground is not perfectly reflecting, however, the resultant magnetic vector will depart from the horizontal, and for any given condition of conductivity its direction, for one sky wave and its reflected field, can be calculated. Figs. 2 and 3 show the "apparent angle of incidence" θ_o computed for the various values of θ, using reflectivity data given by Pedersen[6] for dry soil ($\sigma = 10^{-15}$ electromagnetic units, $\epsilon = 4$) and for wet soil ($\sigma = 5 \times 10^{-14}$ electromagnetic units, $\epsilon = 8$), respectively. On the basis of these data, counting angles of the normal clockwise from *oz* as we look at the loop from the side with the transmitter on our left, and making the proviso that because of the ambiguity in the instantaneous sense of the magnetic force we shall always choose that sense of the normal which gives an angle of less than 180 degrees, it is clear that with a single wave this angle for a null setting should always be in the second quadrant and should never depart from 180 degrees by more than the limiting maximum value of θ_o.

The field resulting from a single downcoming wave has been considered by Verman,[7] with particular regard to the shorter wave lengths, but when there are more sky waves than one the situation is more complicated. The two waves will in general differ in phase, having traversed different paths. The resultants of the two incident reflected systems will combine according to their phase relations and their spatial orientations. Let us consider this case in detail.

We shall assume the existence of two downcoming waves, one with an angle of incidence θ, the other with an angle of incidence θ'. We shall limit our consideration to what happens in the sky-wave furrow, and shall therefore be interested only in that component of each wave which has its magnetic force in the plane of incidence and its electric force at right angles to that plane. For this component, $E_x = E_z = H_y = 0$. Electromagnetic units will be used throughout. Although the waves will in general be elliptically polarized, the projection of each magnetic field ellipse upon the plane of incidence will be a straight line. Two media will be under consideration, air and earth. We shall use subscript zero to designate properties of the former and subscript one of the latter. Letting ϵ represent the specific inductive capacity, μ the permeability, and σ the conductivity, we shall take $\mu_o = \mu_1 = \epsilon_o = 1$, and $\sigma_o = 0$, and consider the surface of the earth as plane in the vicinity of the receiver. c is the velocity of light *in vacuo*.

All the wave trains must satisfy Maxwell's field equations, which will take the form

[6] Pedersen, "Propagation of Radio Waves," Copenhagen, p. 134, (1927).
[7] Verman, Proc. I.R.E., vol. 18, pp. 1396–1429; August, (1930).

$$\left.\begin{aligned} \frac{\partial H_x}{\partial t} &= \frac{\partial E_y}{\partial z} \\ -\frac{\partial H_z}{\partial t} &= \frac{\partial E_y}{\partial x} \\ 4\pi\sigma E_y + \frac{\epsilon}{c^2}\frac{\partial E_y}{\partial t} &= \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}. \end{aligned}\right\} \tag{1}$$

These combine to give the wave equation for E_y,

$$\frac{\epsilon}{c^2}\frac{\partial^2 E_y}{\partial t^2} + 4\pi\sigma\frac{\partial E_y}{\partial t} = \frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial z^2}. \tag{2}$$

This has a solution

$$E_y = a \exp ip\left(t \pm \frac{s}{c}\rho\right),$$

which indicates waves traveling in either sense of $\rho(\equiv lx+nz$, l and n being direction cosines) with a velocity c/s, whence it is clear that s is the index of refraction of the medium. ($s^2=\epsilon-2i\sigma c\lambda$, λ being the wave length in air). We shall consider only the wave in the positive sense of ρ, indicated by the negative sign, and shall consider only real parts as significant. ($i\equiv\sqrt{-1}$.)

Hence for the first incident wave, (subscripts I, R, and T denoting respectively incident, reflected, and transmitted waves),

$$E_{Iy} = a \exp ip\left(t - \frac{x\sin\theta - z\cos\theta}{c}\right),$$

$$H_{Ix} = a\frac{\cos\theta}{c} \exp ip\left(t - \frac{x\sin\theta - z\cos\theta}{c}\right),$$

$$H_{Iz} = a\frac{\sin\theta}{c} \exp ip\left(t - \frac{x\sin\theta - z\cos\theta}{c}\right),$$

H_{Ix} and H_{Iz} being obtained by satisfying (1).

The reflected wave will then be

$$E_{Ry} = f \exp ip\left(t - \frac{x\sin\theta + z\cos\theta}{c}\right),$$

$$H_{Rx} = -f\frac{\cos\theta}{c} \exp ip\left(t - \frac{x\sin\theta + z\cos\theta}{c}\right),$$

$$H_{Rz} = f\frac{\sin\theta}{c} \exp ip\left(t - \frac{x\sin\theta + z\cos\theta}{c}\right),$$

where $f(a)=f_1+if_2$ must be evaluated from boundary conditions.

For the wave transmitted into the earth, the index of refraction $s=s_1+is_2$ will not be unity but will introduce a complex factor into the phase. The introduction of this factor and the necessity of satisfying (1) and the boundary conditions, result in the complexity which appears in the transmitted and reflected amplitudes. We then have (with ϕ=angle of refraction)

$$E_{Ty} = m \exp ip\left(t - s\frac{x \sin \phi - z \cos \phi}{c}\right),$$

$$H_{Tx} = m\frac{s \cos \phi}{c} \exp ip\left(t - s\frac{x \sin \phi - z \cos \phi}{c}\right),$$

$$H_{Tz} = m\frac{s \sin \phi}{c} \exp ip\left(t - s\frac{x \sin \phi - z \cos \phi}{c}\right),$$

where $m(a)=m_1+im_2$ is to be evaluated from boundary conditions.

The boundary conditions are

$$E_{Iy} + E_{Ry} = E_{Ty},$$

$$H_{Ix} + H_{Rx} = H_{Tx},$$

$$H_{Iz} + H_{Rz} = H_{Tz},$$

and substitution of the proper quantities into them yields

$$a + f_1 = m_1$$

$$f_2 = m_2$$

$$(a - f_1) \cos \theta = (m_1 s_1 - m_2 s_2) \cos \phi,$$

$$- f_2 \cos \theta = (m_1 s_2 + m_2 s_1) \cos \phi,$$

whence,

$$f_1 = a\left[\frac{\cos^2 \theta - (s_1^2 + s_2^2) \cos^2 \phi}{\cos^2 \theta + 2s_1 \cos \theta \cos \phi + (s_1^2 + s_2^2) \cos^2 \phi}\right],$$

$$f_2 = m_2 = a\left[\frac{-2s_2 \cos \theta \cos \phi}{\cos^2 \theta + 2s_1 \cos \theta \cos \phi + (s_1^2 + s_2^2) \cos^2 \phi}\right],$$

$$m_1 = a\left[1 - \frac{2s_2 \cos \theta \cos \phi}{\cos^2 \theta + 2s_1 \cos \theta \cos \phi + (s_1^2 + s_2^2) \cos^2 \phi}\right],$$

where,

$$\phi = \sin^{-1}\left(\frac{\sin \theta}{s_1}\right).$$

We have also, from $s^2 = \epsilon - 2i\sigma c\lambda$,

$$s_1^2 = \sqrt{\left(\frac{\epsilon}{2}\right)^2 + (\sigma c\lambda)^2} + \frac{\epsilon}{2},$$

$$s_2^2 = \sqrt{\left(\frac{\epsilon}{2}\right)^2 + (\sigma c\lambda)^2} - \frac{\epsilon}{2}.$$

We can now add the respective colinear components of the incident and reflected waves and obtain the components of the resultant field above the earth caused by the first incident wave. We obtain

$$E_{y1} = [(a+f)\cos d + i(a-f)\sin d]\exp ip\left(t - \frac{x\sin\theta}{c}\right),$$

$$H_{x1} = [(a-f)\cos d + i(a+f)\sin d]\frac{\cos\theta}{c}\exp ip\left(t - \frac{x\sin\theta}{c}\right),$$

$$H_{z1} = [(a+f)\cos d + i(a-f)\sin d]\frac{\sin\theta}{c}\exp ip\left(t - \frac{x\sin\theta}{c}\right),$$

where $d = ip\ (z\cos\theta)/c$. These equations will, of course, satisfy Maxwell's relations, and will therefore represent a possible mode of propagation.

If we turn now to the second incident wave with angle of incidence θ', by analysis similar to the foregoing, and start with an incident wave in which

$$E_{Iy} = b\exp ip\left(t - \frac{x\sin\theta' - z\cos\theta'}{c}\right),$$

we find the resulting field above the earth to be given by

$$E_{y2} = [(b+g)\cos d' + i(b-g)\sin d']\exp ip\left(t - \frac{x\sin\theta'}{c}\right),$$

$$H_{x2} = [(b-g)\cos d' + i(b+g)\sin d']\frac{\cos\theta'}{c}\exp ip\left(t - \frac{x\sin\theta'}{c}\right),$$

$$H_{z2} = [(b+g)\cos d' + i(b-g)\sin d']\frac{\sin\theta'}{c}\exp ip\left(t - \frac{x\sin\theta'}{c}\right),$$

where,

$$g_1 = b\left[\frac{\cos^2\theta' - (s_1^2 + s_2^2)\cos^2\phi'}{\cos^2\theta' + 2s_1\cos\theta'\cos\phi' + (s_1^2 + s_2^2)\cos^2\phi'}\right],$$

$$g_2 = b\left[\frac{-2s_2\cos\theta'\cos\phi'}{\cos^2\theta' + 2s_1\cos\theta'\cos\phi' + (s_1^2 + s_2^2)\cos^2\phi'}\right],$$

$$\phi' = \sin^{-1}\left(\frac{\sin\theta'}{s_1}\right),$$

$$d' = ip\,\frac{z\cos\theta'}{c}.$$

We shall now confine our attention to the magnetic components which will combine in the xz plane. The two components H_{x1} and H_{x2} can evidently be added directly, as can H_{z1} and H_{z2}. The results will be of the form

$$H_x = x_1 \exp iv + x_2 \exp iw,$$

$$H_z = z_1 \exp iv + z_2 \exp iw.$$

Now consider some instant at which $H_x=0$. Then $x_1 \exp iv = -x_2 \exp iw$, whence $x_1=-x_2$ and $v=w$, or, what amounts to the same thing, $x_1=x_2$, and $v=w+\pi$, (for angles less than 2π). But at this particular time H_z will not be equal to zero unless also $-z_1=z_2$, then. In general, such will not be the case, and as a result H_x and H_z will be out of phase, giving a combined effect which will be an ellipse in the xz plane.

There are two particular conditions under which the ellipse will degenerate into a line. One of these is that $\exp iv$ and $\exp iw$ go through zero together, requiring that the two incident reflected field systems be either in phase or in opposite phase. The other possibility is that

$$\begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} = 0,$$

in which case H_x and H_z will go through zero simultaneously for all values of v and w. The vanishing of the determinant means that

$$\frac{x_1}{z_1} = \frac{x_2}{z_2},$$

for which $\theta=\theta'$.

Analysis and the construction of ellipses for the cases of various values of $\Delta\theta = |\theta-\theta'|$ shows the following characteristic properties of the resultant ellipse:

1. The smaller the angle $\Delta\theta$, $(\Delta\theta<\pi/2)$, the greater the number of narrow ellipses near the $\Delta\zeta=0$ straight line and the fewer the narrow ellipses near the $\Delta\zeta=\pi$ straight line, as $\Delta\zeta$ passes through a cycle from 0 to 2π. (ζ_1 and ζ_2 are the respective epochs of the two incident reflected field systems.)

2. The smaller the angle $\Delta\theta$, the smaller the number of broad ellipses in a cycle of $\Delta\zeta$ from 0 to 2π.

3. The smaller the angle $\Delta\theta$, the larger the ratio of the length of

the line for $\Delta\zeta=0$ to the length of the line for $\Delta\zeta=\pi$. Should $\Delta\theta$ become very small, the length of the latter line would approach the difference of the two amplitudes being compounded, and the length of the former their sum.

These properties become very useful in the consideration of experimental results, particularly the generally sharp tilt setting indicating a straight line (or if an ellipse, a narrow one), with only occasionally evidence of a broad elliptic field figure.

Apparatus

The receiver at first utilized an input system as shown in Fig. 4, feeding a neutralized tuned radio-frequency amplifier. All was well shielded, against both interstage reactions and extraneous infiltrations.

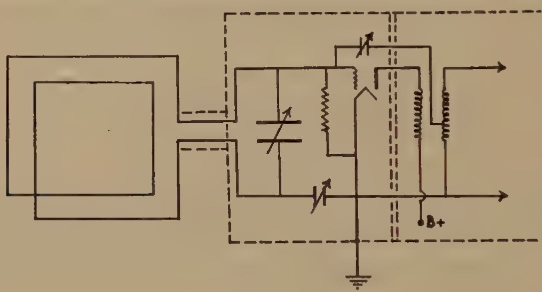


Fig. 4

The incoming signal was made to produce beats of audible frequency with the detector stage, which was kept oscillating, a so-called autodyne arrangement, and the intensity of these beats was used to gauge the amplitude of the radio-frequency electromotive force produced in the coil aerial by the electromagnetic field under observation. In particular, it was desired to find the position of the coil at which the intensity was zero.

UX201A and UX112A tubes were used with a storage battery for filament supply and with dry batteries to furnish anode voltages. This outfit was satisfactory as long as it could be kept *in situ* for studying variations with time, but for portable use the receiver was redesigned.

The new circuit was that of Fig. 5. The utilization of a push-pull radio-frequency amplifier gave a symmetry to the circuit to begin with which greatly simplified the problem of balancing for antenna effect. The balancing arrangement shown in Fig. 5 was a far superior one to what had been used before, especially with screen-grid tubes of the 32 type, which needed no neutralization because of their low mutual capacitance.

In the new circuit the local oscillations were much more tractable, that is, it was found possible to keep them out of the aerial more easily, since their production was delegated to a separate circuit which could be better isolated in the physical layout of the receiver. This was accomplished by the use of a type 1A6 pentagrid tube.

Tubes with two-volt filaments were used throughout, since it was found that these could be operated satisfactorily with two dry cells in series with a voltage dropping resistor. The very lightest type of anode supply batteries available was employed, and all batteries were housed in a small copper box with a cover. This box was grounded to the receiver case, in fact, was set atop it when in operation.

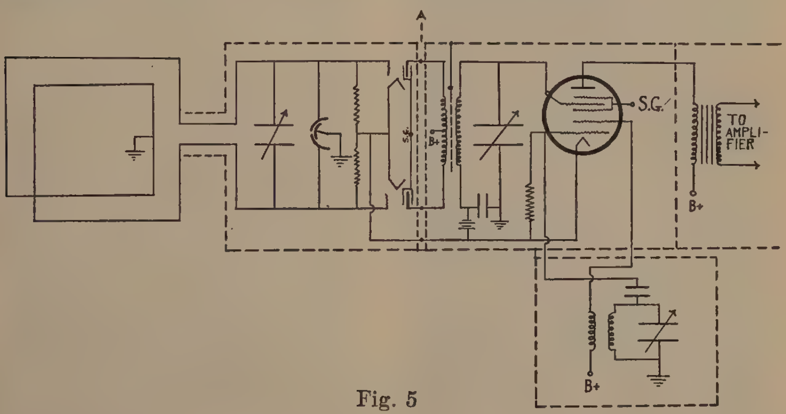


Fig. 5

The coil aerial or "loop" used in observing variations of tilt angle with time was 30 inches square and consisted of twelve turns of bare copper wire spaced 3/8 of an inch apart on a heavy wooden frame. It was designed to rotate about a horizontal axis, as well as about a vertical one, as shown in Fig. 6. A gravitational indicator of tilt angle was devised, consisting of a short plumb line suspended from the center of a circular scale, which made the loop self-leveling, or more precisely made the axes self-orienting with respect to the vertical.

For portable work on location, a very light frame aerial, 25 inches square with thirteen turns of stranded insulated wire spaced 3/8 of an inch apart, also equipped with a gravitational indicator, was used. This appears in Fig. 7, which is a photograph of the apparatus. In use it could be swung on a light wooden tripod base equipped with a wide fork.

In making observations it was usual to place the set upon the ground, and when this was done it made no apparent difference whether or not any further ground connection was made.

The apparatus was considered in balance when the loop could be rotated through 180 degrees, the leads interchanged, and the loop again rotated through 180 degrees, without changing the null reading in any of the four resulting positions. The routine experimental method involved making this test at each place in a location study, taking the reading as the mean obtained, but in studying variations with time it was found impossible to apply the test once variations had started,

Fig. 6

Fig. 7

and therefore it had to be sufficient that the system was balanced at the beginning of observations, after which there was little likelihood of any such throwing out of adjustment as might be effected in moving the apparatus in space. In case any suspicion arose, a local oscillator was kept available to transmit energy from a given direction at a near-by frequency.

Experimental Results

The convention for recording the angle of tilt, α, has been given above.

The resultant of H_{Ix} and H_{Rx} is always positive and the resultant

of H_{Iz} and H_{Rz} is also always positive, (if we choose positive H_{Ix} to start with), since $|H_{Rx}| < |H_{Ix}|$ and $|H_{Rz}| < |H_{Iz}|$, and hence the resultant H in the xz plane will always be in the second quadrant for a single wave. Also by the principles embodied in Figs. 2 and 3 the direction of H would never differ from 180 degrees by more than a certain limiting angle. Whenever it is found to be less than 90 degrees, in fact, whenever it is found outside its limiting angle for a given conductivity and dielectric constant, we would seem to have *prima facie* evidence of the existence of another sky wave (or its equivalent, as will be discussed subsequently) differing both in phase and in angle of incidence.

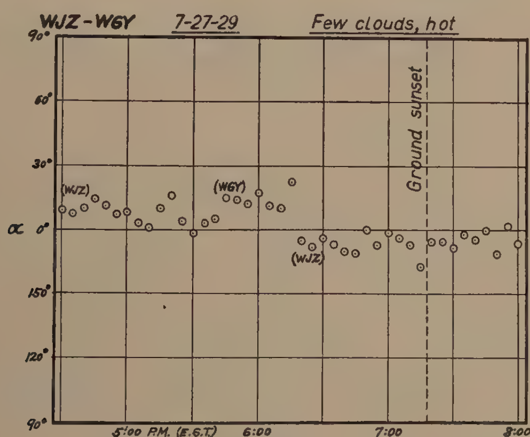


Fig. 8

When a rotation of the sky-wave furrow is mentioned below, the rotation will be defined as the angle between the normal to the plane of the loop and the nearer normal to the xy plane, with the loop set for minimum, when the intersection of the loop's plane with the xz plane is the x axis. This rotation angle will be called β and in recording it the usual procedure was to state in what compass direction the plane of the loop was above the horizontal plane.

Variations with Time

The first part of the investigation consisted of a study of the variation of tilt angle with time. A long series of observations of the diurnal variations of the tilt angle was made, some of the results of which are shown in Figs. 8 to 13. From some of these figures many intervening readings are omitted, to avoid confusion. (The plots are to be regarded as if made on a drum, that is to say, if, for instance, the ordinate extends to 90 degrees at the top of the graph and to 90 degrees at the

bottom of the graph, there is to be understood a continuity between the bottom and the top.) The angle was generally very steady through the day until about two hours before sunset, when variations would usually set in. And whereas the reading through the day would in general have been outside the limiting angle of Fig. 2 or Fig. 3, the drift of the mean reading would then be gradually toward and into this limiting range. Sometimes conditions would steady down again two or three hours after sunset, in which case the angle would as a usual thing remain within the limiting range, but more often the fluctuations would persist throughout the night, respecting the limiting angle only in the mean.

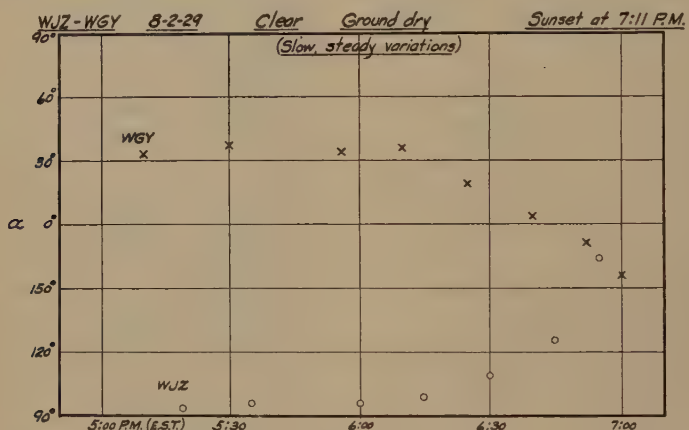


Fig. 9

A few of the records show no lapsing into variations, but rather a continued steadiness of the angle, which very gradually drifted into the limiting range.

On the other hand, on some days, even before the normal sunset period, the variations were so swift and erratic that it was impossible to follow them continuously, but for the most part sufficient approach to instantaneity of readings was attained to gain a fairly continuous record of the sunset phenomena. This is attested by the author's ability to follow and record a number of complete rotations made by the angle.

A fairly large number, though by no means the majority, of the angles observed during the day were less than 90 degrees, and it was for this reason that the possibilities of multiple rays were first investigated. Austin[4] has also observed these first quadrant angles. And even in the case of a great number of the daytime readings above 90 degrees,

it is inconceivable that the ground could be so poor a conductor as to account for them on the basis of a single downcoming wave.

In this connection, it should be noted that in the summer season, when it was necessary that these observations largely be made, the

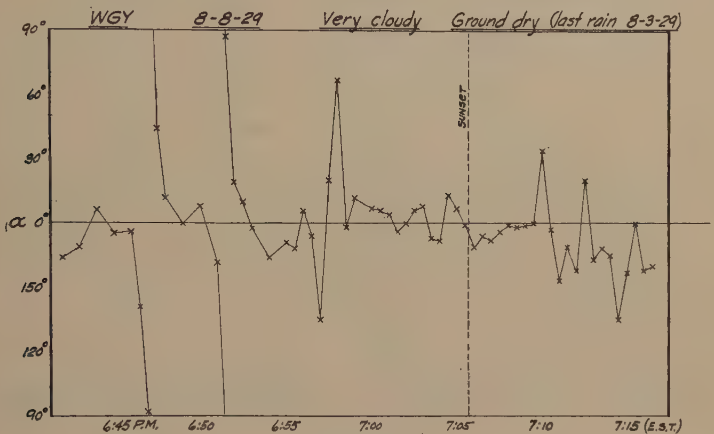


Fig. 10

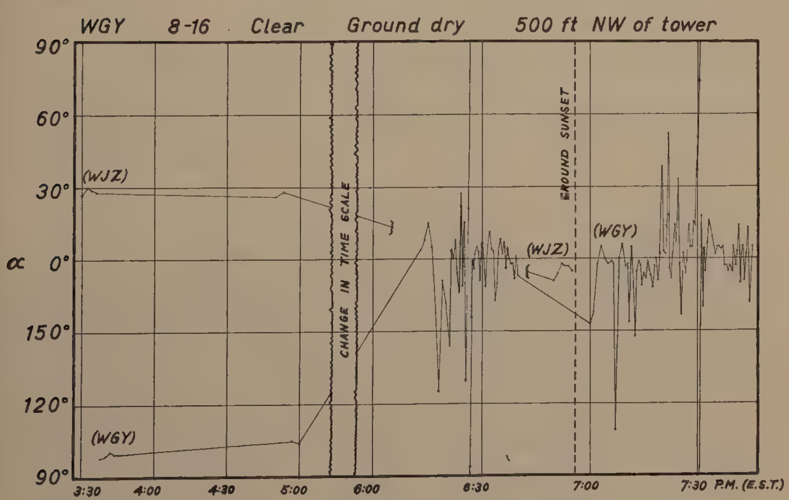


Fig. 11

university grounds become thoroughly dried out. The campus is situated at an altitude of about four hundred feet above the city, so that the drainage of moisture is very effective. The condition of the soil particularly in the advanced stages of a very dry summer assumes a high degree of desiccation. For example, at the time of the particular

observations shown covering a perpendicular transit of the cinder track, the conductivity of the surrounding field must have been well below 10^{-15} electromagnetic units.

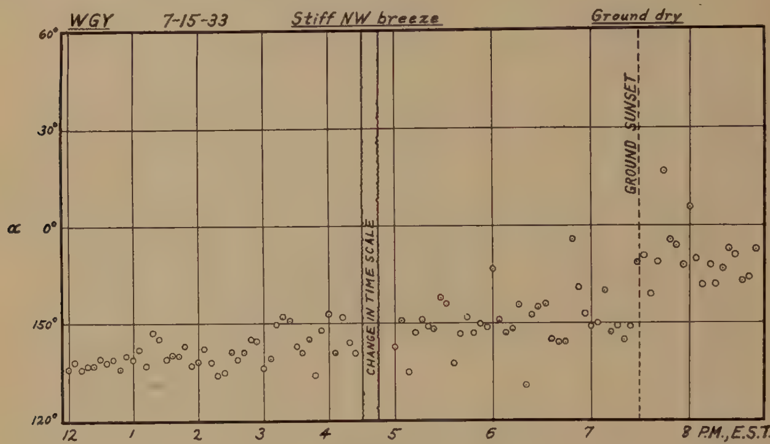


Fig. 12

The occurrence of rain was found to result in a very definite decrease in any angles below 90 degrees and an increase in those above 90 degrees. In morning observations, dew was often found on the

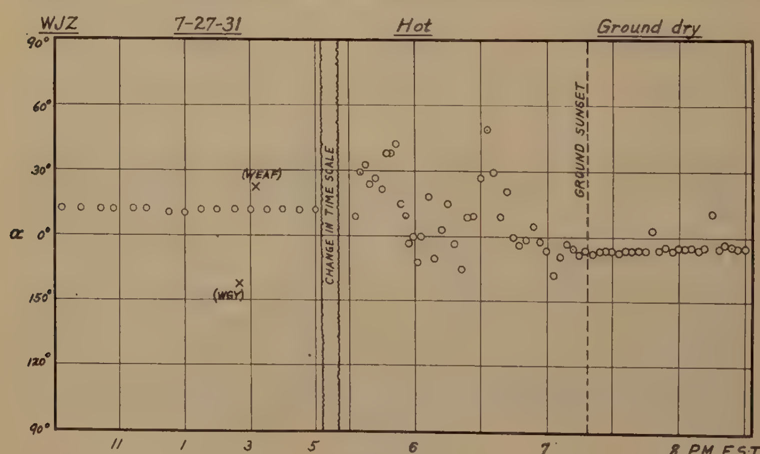


Fig. 13

grass and strangely enough in its presence any deviations from zero degrees were usually negligible. This would seem to be attributable to a layer of wetness on the ground beneath, rather than to the discrete globules on the grass. Consideration was given to the possibility that

the clustering of evening angles about zero degrees was caused by dew formation. This may be a contributing factor, but the beginning of the phenomena involved generally precedes the time incidence of the reaching of the dew point, and definitely has occurred on evenings when no dew was formed.

Uusually the minimum found was apparently a zero, or nearly so, within the limits of background noise. There were a few occasions, however, when the magnetic field ellipse in the plane of incidence was very obviously a broad ellipse and no satisfactory setting could be made. It is hoped to obtain further data bearing on this point by an improvement of the method which was devised some time ago but the following up of which circumstances have prevented. The results so far obtained indicate that the ellipse is most often a thin one.

The observation of a normal quiescent period through a large part of the daylight hours laid open the valuable possibility of the study of space variations of the tilt angle. There were, however, many days on which the sky-wave intensity was so small as to be negligible, and at these times observations were impossible. It seems likely that on these days ionization and therefore absorption reached a particularly high value.

Records of the magnetic elements very kindly placed at the author's disposal by the magnetic observatory at Agincourt seem to show sufficient correlation with tilt angle data to enable us to say that periods of pronounced magnetic variations are apt to be periods of unusual fluctuations in tilt angle.

Variations in Space

(1) Case of Buried Horizontal Conductor

The Cornell University Upper Alumni Athletic Field provided a very favorable setting for the investigation of this case. A pair of steam pipes 1 foot, 9 inches apart, one of 8.625 inches outside diameter, 7.981 inches inside diameter, the other of 6.625 inches outside diameter, 6.065 inches inside diameter, and buried at a mean depth, center line, of 3.7 feet and 3.8 feet, respectively, cross the field in the direction N9°E–S9°W. The larger pipe is on the western side. In recording data "the pipe" is considered to be the central axis of the system. The bearing of WGY here is N87°E, so that its plane of incidence is cut nearly perpendicularly by the pipe. The bearing of WJZ is S21°E and hence its ground wave has a very large component Poynting vector along the axis of the pipe. WEAF with a bearing of S35°E provides an angle with the pipe between the two others.

It was found that the field perturbations due to the pipe line were

greater the greater the component of the ground-wave Poynting vector in the axial direction of the pipe line. For a station having a large component (i.e., for WJZ and in lesser degree for WEAF) the tilt of H is toward the transmitter on the farther side of the pipe and away from it on the side nearer the transmitter, and the plane of the loop is horizontal above the line, as in Fig. 14.

The effect is apparently a ground-wave phenomenon, not only on the consideration of intensity of the effect, but also because there is in the closer side positions a rotation of the sky-wave furrow such as could only be brought about by a substantial tangential magnetic field co-

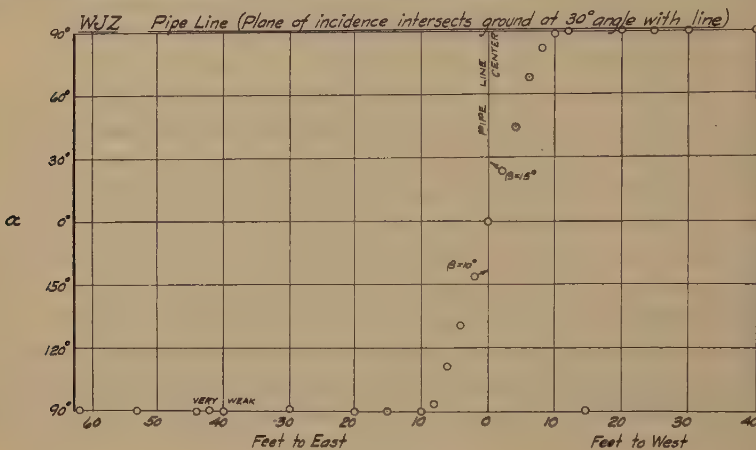


Fig. 14

axial with the pipe and varying with the same frequency as that of the incoming impulses. This magnetic field is doubtless caused by currents in the pipe due to the differences of potential radially from the station produced by the waves.

The situation we have been describing, however, is that obtaining for a very substantial component of the ground-wave Poynting vector along the pipe and therefore for large currents along it. When the currents are much reduced, as for instance when they are due to a sky-wave electric force directed along the pipe, the result is a little more complicated, and is most simply explained by considering the field about a current bearing straight conductor as superposed upon a steady magnetic field. And we can consider that we have such a conductor of conductivity $\Delta\sigma = \sigma_2 - \sigma_1$ (σ_1 being the conductivity of the earth and σ_2 that of the pipe) which exists in the neighborhood of the earth's surface. Consider, for instance, Fig. 15, drawn in the xz plane. Suppose we have the steady (in direction and amplitude) magnetic force $H = H_o \sin pt$ per-

vading the space just above the earth's surface S and produced by a wave coming down from the left and its reflected wave. Obviously the electric vector in the field will be into the paper at an instant when H is as shown, and the current, which will probably be not much out of phase with it, will produce the tangential magnetic field h as indicated about C.

Now suppose we start out at the right and move toward the left, observing the direction of the resultant magnetic field as we go along the surface of the earth. As we come into the sphere, or rather cylinder, of influence of h, the vector will gradually bend down from the undisturbed position indicated toward the horizontal as h becomes stronger and stronger. If h becomes strong enough while we are still to

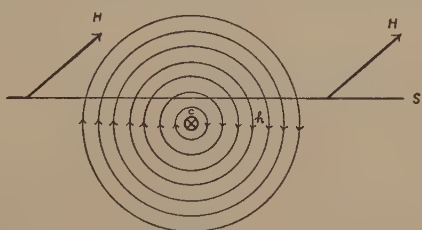


Fig. 15

the right of C, the vector may even be bent down beyond the horizontal. Directly above C, however, h has become horizontal, so that the resultant has come up again from its extreme dip, and from then on it rises gradually toward its original orientation. It should be noted, though, that if h is sufficiently strong in the case of the angle we have here chosen for H, the resultant may be raised (for an interval) beyond the original angle when it gets to the outer parts of h at the left, because of the angle of the h vectors out there.

Maxwell has shown the result[8] of a superposition of this general nature.

Figs. 16 and 17 illustrate observations of the type mentioned. The effect is still more marked in the data taken about the cinder running track, with the modifications which will be discussed in that section. The pipe of Figs. 16 and 17 is nearly, but not quite, normal to WGY's plane of incidence. The two figures show the effect of the external angle of tilt on the symmetry of the resultant field pattern. The fact that the external angle was zero in Fig. 17 seems to indicate that the tangential field is due largely to the ground wave, despite the small

[8] Maxwell, "Electricity and Magnetism," Third Edition, vol. II, Fig. 17, p. 117, (1904).

radial current component. If this is true also of Fig. 16, then in the latter case there seems to have been a larger ratio of sky-wave to ground-wave intensities.

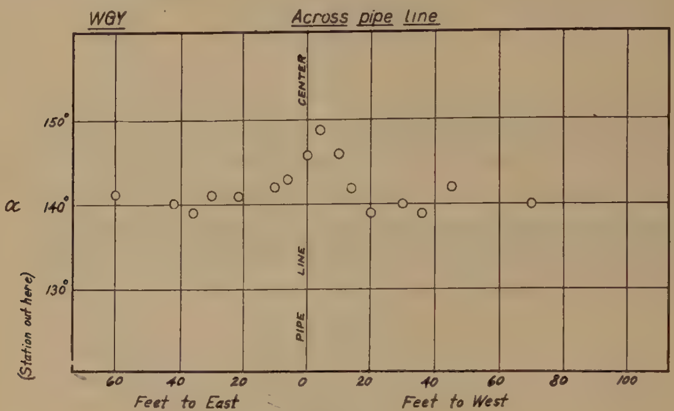


Fig. 16

In the interpretation of the results of horizontal currents due to the ground wave, we must remember that the perturbation field h effective in the observations is the projection on the plane of incidence

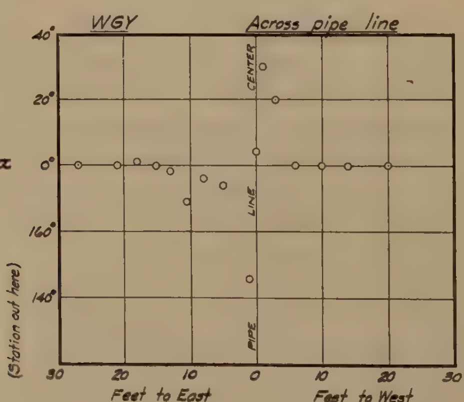


Fig. 17

of a family of cylinders about the conductor (in the case of a simple cylindrical conductor).

Observations near wires on the ground and suspended above the ground yielded the same general results as in the case of the buried pipe, though effects were of smaller magnitude.

(2) Case of Running Track

Another form of conductor was afforded by the oval cinder track in one section of Upper Alumni Field. Hence it seemed of interest to examine a wave field in its neighborhood. The straightaways run S82°E, nearly toward WGY, and a point can be found at either end of the oval where the curved part will pierce WGY's plane of incidence normally. Readings were taken crossing the track along the intersection with the plane of incidence at such a point, and also crossing the straightaway perpendicularly. Examples of these observations are

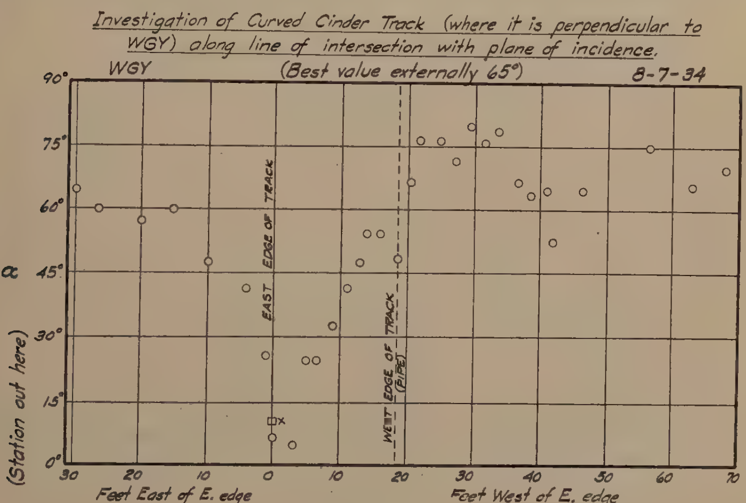


Fig. 18

shown in Figs. 18 and 19. In both of these plots it will be seen that the results can be explained by superposition upon the external field of a tangential perturbation field h about the track's section. Evidently the track, from frequent rolling and watering, has become a closely packed vein of much greater conductivity under usual circumstances than the surrounding field. Smith-Rose[9] has found in some laboratory measurements at high frequencies that packing soil tightly together increased its conductivity.

Consider Fig. 18. Inasmuch as we have seen that the external field H in the plane of incidence, having as noted a tilt angle of 65°, cannot be produced by a single incident wave, let us take it as a vector in the first quadrant and consider it broken up into two components, say H_1 and H_2, each of which could be caused by a single incident wave from

[9] Smith-Rose, *Proc. Roy. Soc.*, ser. A vol. 140, p. 359; May, (1933).

the left, as illustrated in Fig. 20. (The actual components could be more than two.)

Now, the magnetic vector H_1 would be associated with an electric vector into the paper, supposing the paper to represent the plane of incidence, and the current produced in any conductor into the paper

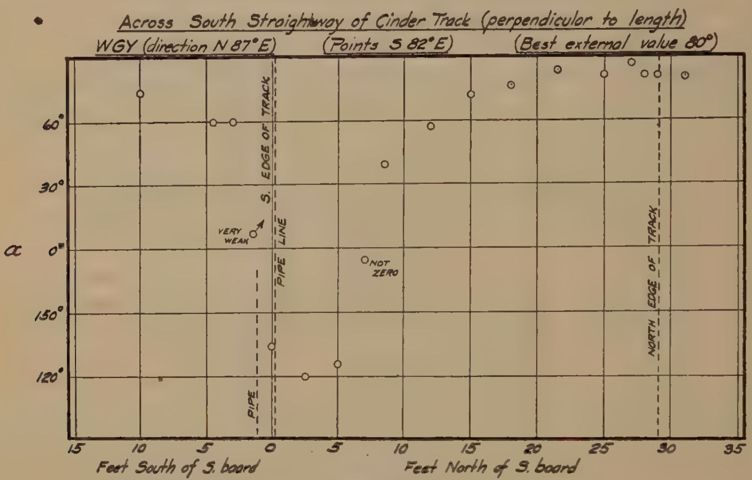


Fig. 19

would have a tangential magnetic field as indicated in Fig. 20 (a) by the circular arrows. H_2, on the other hand, would be associated with an electric force out of the paper and the magnetic field produced by a current in that direction would be counterclockwise as in Fig. 20 (b). Examination of Fig. 18 will show that if we consider H as a vector

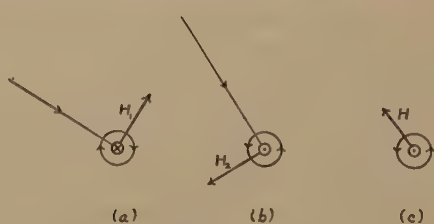


Fig. 20

tilted toward the transmitter, the perturbation field necessary to produce the observed position of maximum dip must be counterclockwise as in Fig. 20 (c), and therefore the electric force must be toward us out of the paper. This would seem to indicate that of the two component vectors, H_2, the one in the wave of smaller angle of incidence would be the larger. All this is on the assumption that H and the perturbation field are in phase.

Fig. 21 shows the results of observations made along the center line of the surface of the running track at the western end. The numbers represent the tilt angles at the positions where they appear. There is apparently a complicated superposition upon the undisturbed field of the fields due to currents set up by both ground wave and sky wave.

(3) Case of Vertical Conductors

The effect of a vertical metal lamppost upon the surrounding electromagnetic field was very striking. Tall posts on the campus of the College of Agriculture about twenty feet in height, were very well suited to this study. The phenomena are best described by specific ref-

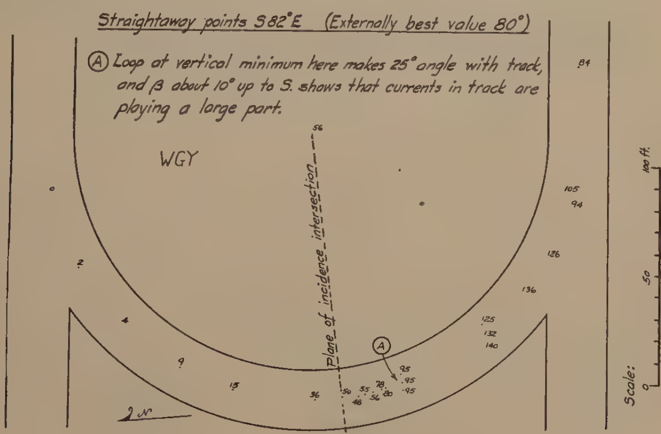


Fig. 21

erence to WGY, bearing N87°E. Observations on WGY showed on approaching the post from the south a decrease in the tilt angle, which externally was 125 degrees, finally to 50 degrees close by the post; on approaching from the north an increase in the tilt angle, reaching 156 degrees near the post. These results are plotted in Fig. 22. But there are other aspects to the resulting field which cannot well be represented in a figure. Along an east-west line, readings tended to remain in the undisturbed direction of the external field, subject only to a few minor irregularities attributed to differences in the nature of the ground. In other directions close to the post there was a slight rotation (less than five degrees) of the sky-wave furrow. All these phenomena apparently indicate a disturbing magnetic field coaxial with the post. The magnitude of this field points to currents in the post doubtless attributable to vertical electric force in the direct wave.

The results with metal posts led to trials on wooden telephone poles. No variations were detected near a dry pole, but with a wet one

there was a change in the undisturbed tilt angle in opposite directions on the two sides, though of very small magnitude compared to the case of the metal post, and with no trace of rotation of the sky-wave furrow.

The inference regarding the trunk of a tree is obvious. In practice, with the rather short trees available, the ramifications overhead probably exerted a complicating influence, for the results showed less system than was expected.

It is apparent that in the neighborhood of any such vertical conductors, especially metal ones, and particularly if there were any ap-

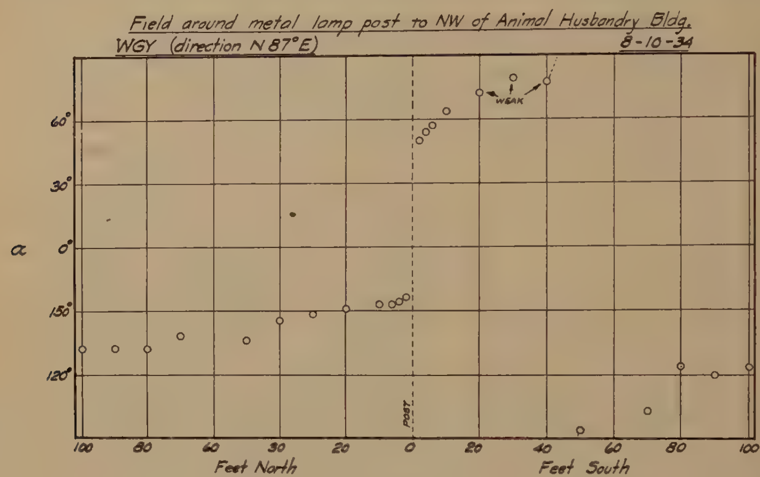


Fig. 22

proach to resonance conditions, there would be in bearings taken in the ordinary way considerable errors; i.e., rotations of the sky-wave furrow about the z axis.

(4) Case of Deep, Wide Gorges

The author first conceived the idea of observing at the edge of a gorge with the notion that here a true reading of angle of incidence might be obtainable because the gorge might act, at least approximately, as an absorbing black body for the reflected wave. It was therefore decided to investigate the possibility. As a matter of fact, it was in order to aid in interpreting the results obtained near gorges that some of the foregoing studies were undertaken.

Deep, wide gorges are characteristic parts of the terrain in and around Ithaca, so that it was possible to observe the phenomena connected with gorges making various angles with the plane of incidence.

(The gorges used had depths of the order of 200 feet, and widths of 200 to 300 feet or more.) To find a gorge which would be accurately normal to the plane of incidence, however, required considerable scouting of locations with a compass and contour maps. It will be evident that the gorge must be as accurately as possible normal to the plane of incidence, because compared to the magnitude of the ground wave in the daytime the sky-wave amplitude is practically a second order quantity, and it would not take much of a component of ground-wave current to mask completely any sky-wave effects. The data seem to indicate that elimination of such ground-wave influence was indeed achieved, except where noted.

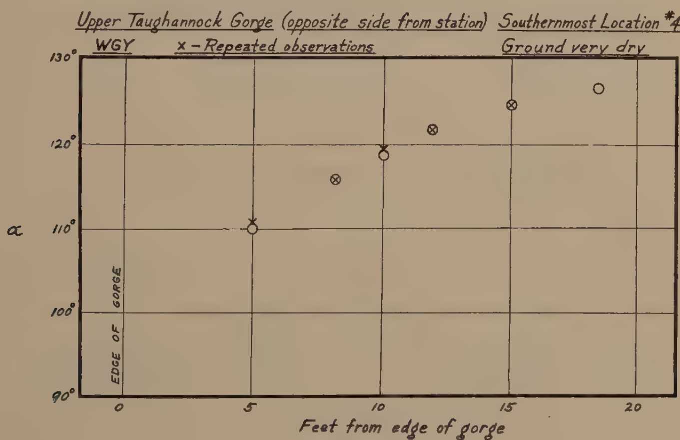


Fig. 23

When it is realized that it was also necessary to have a sufficiently extended stretch of level ground back from the edge of the gorge, as free as possible from trees, the magnitude of the initial problem of seeking out suitable locations will be appreciated.

Some of the observations obtained are plotted in Figs. 23 to 34, inclusive. It is found that the results are in general amenable to interpretation on the basis of the combination of the undisturbed external field with a tangential perturbation field. In the case of a gorge, however, a modification of the conductor case is necessary. We must here subtract from the external field the tangential field which would have been contributed by currents in the missing section had it been present. In other words, we add the field about an "anticonductor" of conductivity $-\Delta\sigma$ (or $\sigma_0-\sigma_1$, where σ_0 is the conductivity of the air and σ_1 the conductivity of the earth). To get at this matter exactly requires some modification of the simple process as expressed, even

assuming a circular cylindrical cross section, since the removed portion of conductor is a hemicylindrical piece—not the complete cylinder—and also the high-frequency current distribution in a conductor

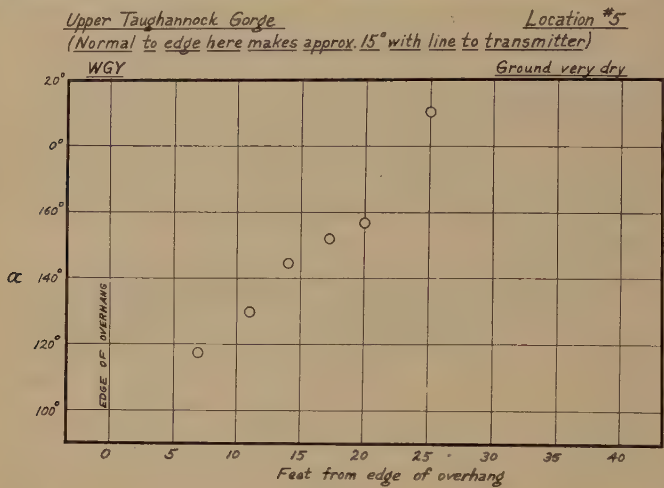


Fig. 24

is not independent of the presence of other parts of the conductor. The author has in preparation a theoretical treatment of this case.

The observations on the side of the gorge opposite the transmitter

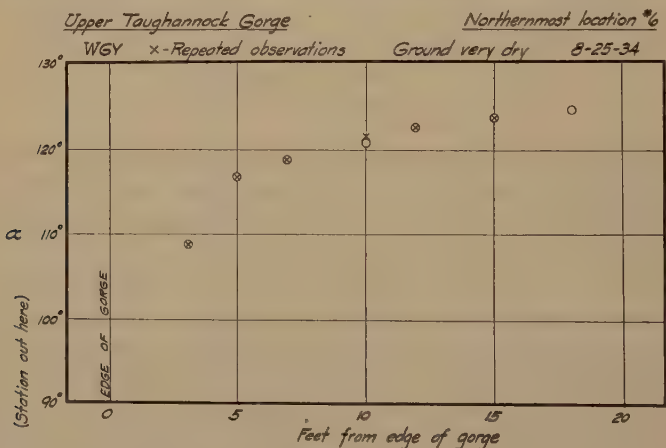


Fig. 25

can be classified as of three types: (1) external magnetic vector in first quadrant and perturbation field due to sky wave; (2) external magnetic vector in second quadrant and perturbation field due to sky wave; (3) perturbation field due to ground wave.

There is no point in resolving type (3) as we did (2) from (1), because the level stretches near gorge edges were comparatively short and it is evident that one would have to move away a considerable distance before the ground-wave perturbation field would be sufficiently reduced in amplitude to be comparable with the sky-wave field.

Type (1) is shown in Figs. 28 and 30; type (2) in Figs. 23, 25, and 32; and type (3) in Figs. 26, 29, and 31.

Fig. 27 is an example of type (3), as well as an illustration of the effect of a group of three conductors in the waist-high metal rail fence safeguarding sightseers from falling into the gorge.

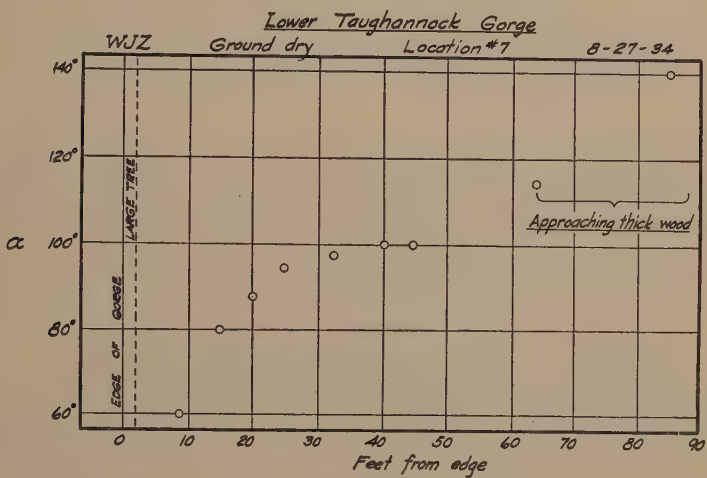


Fig. 26

Fig. 24 seems to be a composite due to the combination of local conditions, ground wave, and sky wave. The last point obtained, at twenty-five feet from the overhanging edge, was directly adjacent to a large tree, and to push on farther would have necessitated climbing a bank among other trees. The large angles observed at the other points were attributed to the fact that the mean gorge side face made an angle of about 75 degrees with the plane of incidence. The edge jutted out seven or eight feet at the top in a precariously shallow overhanging shale shelf, upon which it was not deemed wise to venture far with the apparatus.

The cause for the increase in angle at the larger distances in Fig. 26 may have been the approach to a thick wood. In fact, the farthest point taken was at the very entrance to a road into the wood.

In Fig. 28 is seen again the effect of the metal fence, this time because of sky-wave currents. These data in the neighborhood of twenty

feet from the edge seem to show the effect of some local irregularity which was not otherwise evident, although this may be an example of return from a slight overcorrection of dip as previously discussed, in case (1).

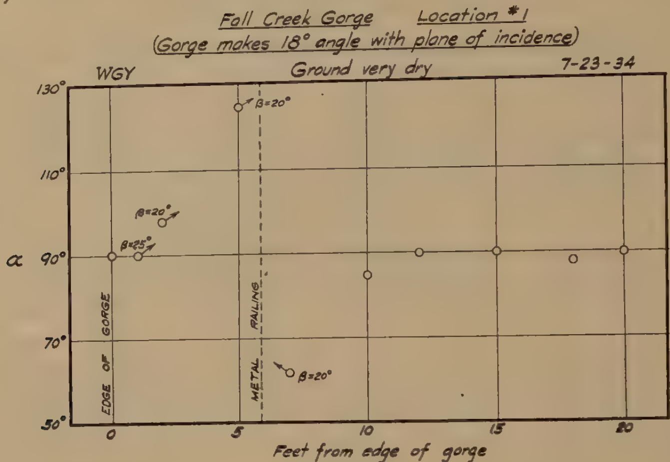


Fig. 27

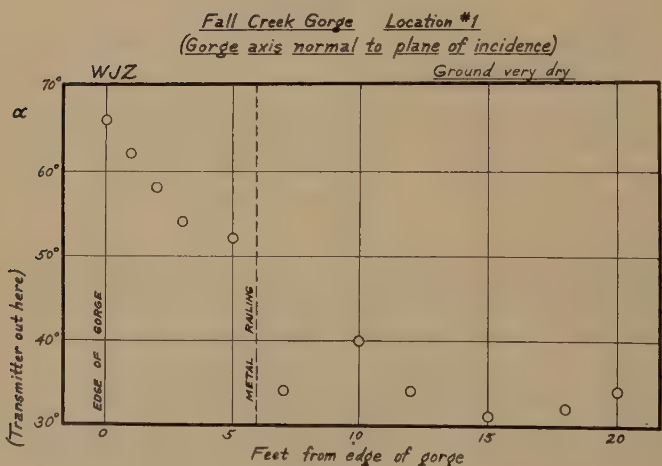


Fig. 28

Location 2B at Fall Creek Gorge was a very fortunate one, equipped with a ten-foot projection a few feet wide. It was at first feared that this might distort the field, but it apparently did not. Data obtained there in general showed a distinct falling off in rate of change of tilt angle just beyond the regular edge, a very valuable indication which will be mentioned under "applications."

In some of the figures mentioned the rotation of the sky-wave furrow about the x axis as a result of ground-wave currents is shown by arrows indicating the approximate direction of the plane of the loop, which is ordinarily horizontal for $\beta=0$.

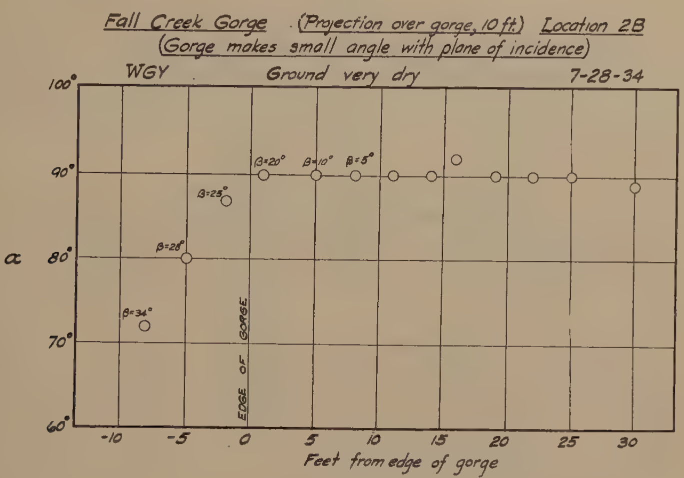


Fig. 29

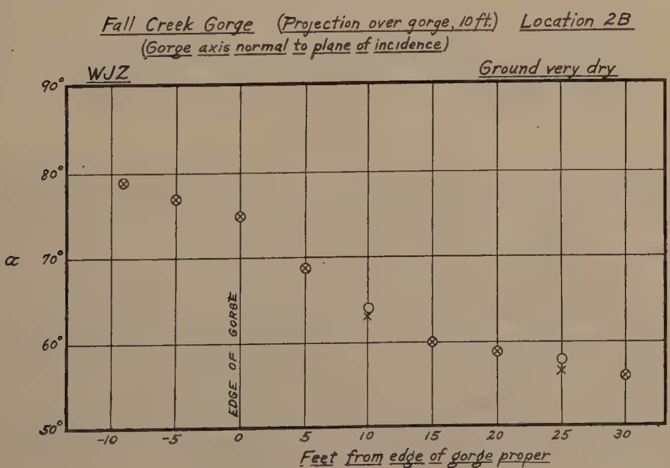


Fig. 30

Observations on the side of a gorge nearer the transmitter, plotted in Figs. 33 and 34, show the same general mode of variation as the ones taken on the opposite side. Here the situation was complicated beyond about thirty feet from the edge by a further ten-foot incline of nearly 45 degrees up from the shelf at the gorge edge.

The gorge bottoms were so thick with trees or irregular in other ways that it was not believed any significant readings could be obtained there.

(5) Effect of Change of Reflecting Surface

In the light of the results of the experiments with conductors and with gorges, it is of interest to note the results of seeking a surface—a very broad expanse—of very different conductivity, while conditions remain steady.

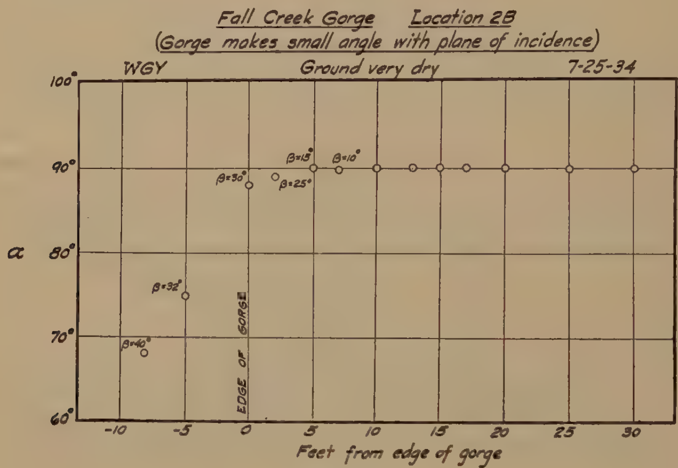


Fig. 31

The first attempts to get readings above Cayuga Lake were made at the white lighthouse off Stewart Park, the eastern one of two widely separated lights. This one stands at the end of a long breakwater, a location which proved very unsatisfactory for several reasons. The breakwater and lighthouse platform are of concrete reinforced with metal and there are several metallic conductors (wires and pipes) leading to the lighthouse. Considerable annoyance was experienced because of the fact that the light continually goes on and off automatically with a period of about two seconds. The resulting clicks in the receiver phones make settings very difficult.

Recourse was had, therefore, to a rowboat. A day was selected when the lake was fairly calm and the tilt angle for WGY and WJZ in different quadrants. Observations were first made on Alumni Field, and the apparatus was then taken out on the lake.

It was in taking data thus under quasi-nautical circumstances that the full value of the gravitational tilt indicator was appreciated, for

there was no doubt as to its indicating the correct angle with respect to the vertical at all times. The apparatus was mounted on the after seat of the boat which was kept headed at right angles to the plane of incidence. As the boat did not displace much water, the "gorge" created would be of small dimensions.

The mean values of the observations on the field were for WGY, 139 degrees; for WJZ, 32 degrees. On the lake they were for WGY, 179 degrees; for WJZ, one degree.

Readings taken over salt water just inside the mouth of the Mer-

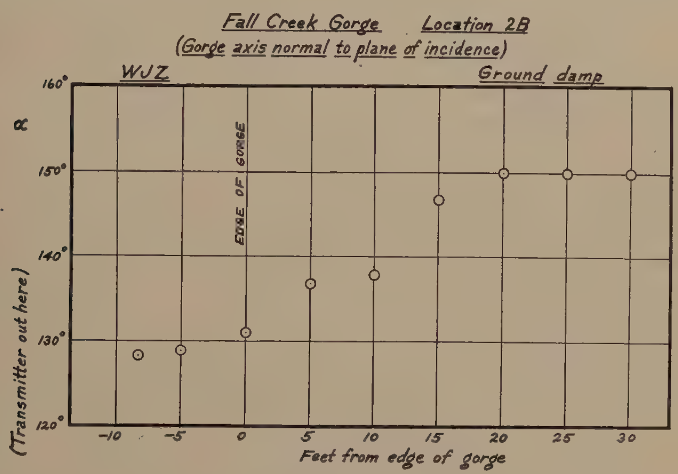


Fig. 32

rimack River at Salisbury Beach, Massachusetts, have shown negligible variations from zero.

General Observations

A few phenomena of a general nature were observed which have not been discussed in the foregoing parts.

One of these was an occasional complete rotation of the observed magnetic vector, usually in the sunset period. This would seem to be explainable, in fact to be called for, on the multiple wave hypothesis as presented here. When one of the wave components is changing its length of path continuously relative to that of the other, as when the ionospheric height is changing, there will be a resulting change of relative phase which will cause the magnetic field ellipse in the plane of incidence to rotate.

Over the series of observations, also, the distinct impression was gained that the intensity of the magnetic vector was less when found

in the first quadrant than when found in the second quadrant. Sometimes, in fact, a range of silence of as much as twenty-five or thirty degrees in breadth occurred in the former case.

The tendency for WJZ to have a tilt angle considerably smaller than WGY is very noticeable in the data.

Applications

Several applications of the foregoing methods immediately suggest themselves.

1. We have a very sensitive method of locating underground veins of some size having different conductivity from—either greater or less than—their surroundings. It is of interest to note that we had for a long time believed (though we had never had occasion to make use of the apparent fact) that there was another pipe line forty-two feet eastward of and parallel to the main steam line previously mentioned. Tilt observations, however, showed absolutely no variations upon crossing that line, and a consultation of the records of the buildings and grounds department disclosed that there was no pipe there. The straightness of the line of dead grass (made therefore by pedestrians) of similar appearance to that above the steam line, and its parallelism to the latter, might be matters to occupy the psychologist.

2. An indication of the phase of the sky wave relative to the ground wave could be obtained. But here again one is led to the necessity for a study of the actual ellipse forms involved in the resultant field.

3. Interesting considerations in the placing of direction finding apparatus have been here developed.

4. If the gorge can be used to take out the reflected wave, a direct observation of the angle of incidence may be made under certain conditions. This matter the author intends to treat more fully in a subsequent paper.

5. Observations above a gorge or any nonconductor of considerable size provide a method of finding the relative intensity of the reflected wave, and therefore the conductivity of the ground.

It seems reasonable from previous considerations that in the case of a very wide and deep nonconductor like the gorge a broad area would be found above it in which the tilt angle retained its maximum approach to ninety degrees. The data with a projecting edge lend credence to the belief in the existence of such an area, beginning not far from the edge, as indicated by the decrease in rate of change of tilt beyond the edge. That such an area would be characterized by incident wave field, with the reflected wave practically eliminated, ap-

pears reasonable on the following considerations. We may consider the reflected wave as caused by the ground currents set up by the incident wave. Therefore, the limit approached over a section of perfect conductor would be a reflected wave exactly equal to the incident

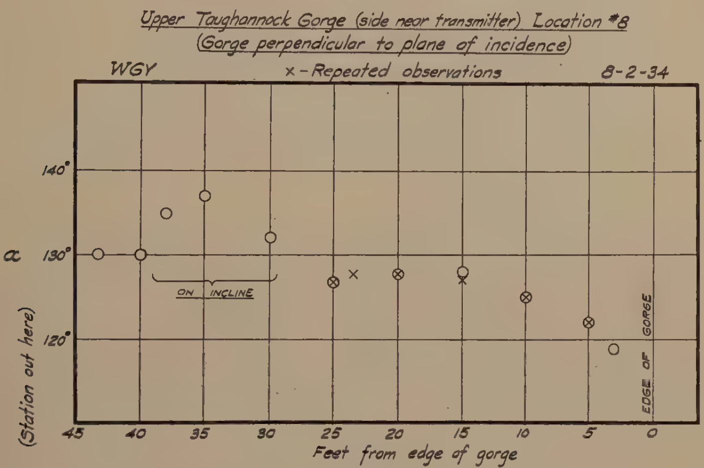


Fig. 33

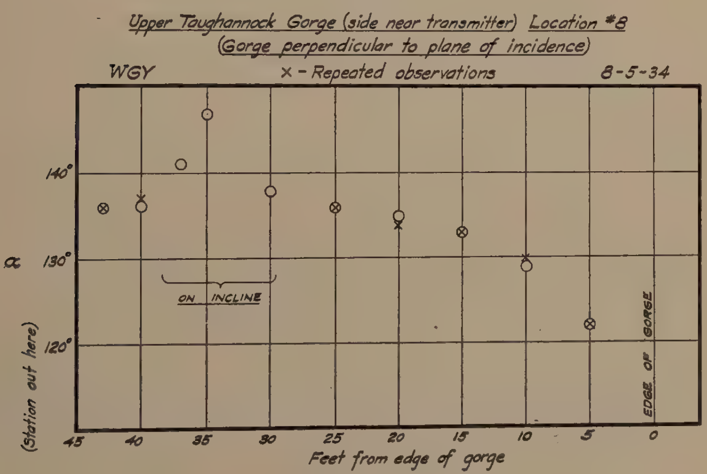


Fig. 34

wave, if we were sure the significant ground currents were due only to the incident wave. By the same reasoning, the limit approached above a section of perfect insulator would be removal of the reflected wave, and this should be closely approximated above a deep, wide gorge, particularly in view of its likely action as a black body for radiation falling into it.

Now, to get a quantitative value for the coefficient of reflection, it would be necessary to take a reading of the tilt angle for the external undisturbed field and one for the maximum approach to ninety degrees. It would then be a simple matter to find the ratio of reflected magnetic vector to incident magnetic vector.

For routine measurement of conductivity it might be possible even to construct an artificial nonconducting "gorge" for the convenient making of the necessary readings. In order always to be sure to have a downcoming wave, it would also be possible to produce one's own downcoming wave by placing a transmitter on some near-by eminence, such as a tree, tower, or kite.

Possible Origins of Multiple Sky Waves

As we have seen, it seems impossible to account for the angles observed outside the limiting values by any theory comprehending a single sky wave, unless one is willing to admit the possibility of a horizontally polarized ground-wave component. We know that there should be a Faraday rotation of the plane of polarization of the direct wave because of the presence of the earth's magnetic field, but the usual assumption is that the vertical component of magnetic force so formed is damped out immediately because of the conductivity of the earth's surface. It is conceivable, however, that the earth might under certain conditions be a poor enough conductor to allow a small component to be built up. In the daytime at the distances used the downcoming intensity is practically a second order quantity compared to that of the direct wave, so that a second order effect in rotation of the plane of polarization of the latter could conceivably produce a vertical magnetic force comparable with the magnetic force in the field produced by the sky wave. As noted in the experimental discussion, the tilt angle of WJZ tends to be less than that of WGY, which at least can be said to be a result in a direction not at variance with such a hypothesis, since the line between Bound Brook and Ithaca would coincide with a much larger horizontal component of the earth's magnetic intensity than would the line between Schenectady and Ithaca, although both would be small components of the total magnetic intensity because of the large dip. If the observed differences in tilt angle should be due to rotation of the plane of polarization of the ground wave, arid land near the equator should show a marked effect, with stations not too far distant to the north and south giving a tilt angle seldom departing far in the daytime from ninety degrees.

Now let us turn to the possibilities which might give rise to two or more sky waves.

First, there is the suggestion advanced by Merritt[10] that there may be a wave of low trajectory received at a very glancing angle. The action of the horizontally polarized component of such a wave upon the receiver would be practically the same as that of the horizontally polarized ground wave discussed above, and therefore must be counted as a possible cause of the observed phenomena. The principal difference in its propagation would be that, since there is no conducting surface along most of its path, it would seem that there would be nothing to limit its plane of polarization from rotating into any position.

Such a wave could easily be produced if there were a sufficient number of ions not far above the earth's surface. The observations reported by Colwell and Friend[11] and by Watson Watt, Bainbridge-Bell, Wilkins, and Bowen,[12] indicating the presence of one or more very low regions of high ionization, appear to bear out Merritt's prediction. The absence of evidence of such regions from the records of conductivity made on the Stevens-Anderson stratosphere trip[13] of November 11, 1935, and from the data of previous investigators may mean their seasonal, and possibly longer period secular, occurrence.

Another possible mode of production of a second indirect wave would be by magnetoionic double refraction, of which the investigation is now far advanced.

Lorentz[14] developed the theory for an electromagnetic wave traveling through a system of ions or electrons under the influence of an applied magnetic field. The theory was applied to the action of the ionosphere and the earth's magnetic field upon radio waves by Eccles,[15] Larmor,[16] Pedersen,[17] Nichols and Schelleng,[18] Appleton,[19,20,21,22] and others. For a brief view of the historical development of the subject and a suggestive bibliography up to that time, the reader is referred to an article by Kirby, Berkner, and Stuart.[23]

[10] Merritt, "Optics of Radio Transmission," *Jour. Opt. Soc. Amer.*, vol. 21, p. 94; February, (1931); Proc. I.R.E., vol. 22, pp. 29–39; January, (1932).

[11] Colwell and Friend, *Nature*, vol. 137, p. 782; May 9, (1936).

[12] Watson Watt, Bainbridge-Bell, Wilkins, and Bowen, *Nature*, vol. 137, p. 866; May 23, (1936).

[13] Gish and Sherman, *Nat. Geog. Soc. Stratosphere Series No. 2*, Washington, p. 105, (1936).

[14] Lorentz, "Theory of Electrons," Leipzig, (1909).

[15] Eccles, *Proc. Roy. Soc.*, ser. A, vol. 87, p. 79; August, (1912).

[16] Larmor, *Phil. Mag.*, vol. 48, p. 1025; December, (1924).

[17] Pedersen, "Propagation of Radio Waves," Copenhagen, (1927).

[18] Nichols and Schelleng, *Bell Sys. Tech. Jour.*, vol. 4, p. 215; April, (1925).

[19] Appleton, *Proc. Phys. Soc.*, vol. 37, part II, p. 16D; February, (1925).

[20] Appleton, *Proc. U.R.S.I.*, part I, p. 2, (1927).

[21] Appleton, *Jour. I.E.E.* (London), vol. 71, p. 642; October, (1932).

[22] Appleton and Naismith, *Proc. Roy. Soc.*, ser. A, vol. 137, p. 36; July, (1932).

[23] Kirby, Berkner, and Stuart, Proc. I.R.E., vol. 22, pp. 481–521; April, (1934).

The fact of the bending back down to earth of radio waves sent out into space depends upon an increase in velocity of propagation with height, that is, a decrease in the index of refraction with height. Larmor[16] showed that a change in specific inductive capacity must be involved, since to accomplish the necessary bending by increase in conductivity alone would introduce so much absorption as to be prohibitive of the return of any of the wave to the earth. Expressions for the amount of this change in dielectric constant have been given by Pedersen.[17]

Whether a continuous or a discontinuous increase in ionization is assumed as height increases in the "layer," the wave directions at any two levels are generally taken[22,23,24,25,26] to be related by Snell's law, assuming the ionization a function of height alone. If we then assume that there can be no reflection from such a medium, the question of what takes place physically at the top of the path in zero base echo sounding, where the wave is sent vertically upward, becomes subject to considerable discussion, for in such a case $n=0$ appears as the condition to be satisfied.[22,23] Epstein[27] has indicated, however, that continuity or discontinuity of n is immaterial so far as the possibility of total reflection is concerned, but that modifications are necessary in Snell's relation, as we should have expected in view of the wave lengths involved.

Electric waves are found returning from two general regions of the ionosphere (above the newly projected regions), known as the E layer, at a height of 100 to 120 kilometers, and the F layer or region, at 200 kilometers and above. An expression for the index of refraction n has been worked out by Appleton and his colleagues[22,28] in which the frequency is a parameter, as well as the number of electrons per cubic centimeter. The higher frequencies are returned from the higher levels, and a value for the number of ions is often found from the critical frequency, or frequency which will just penetrate completely a certain layer.

It is not considered that the E and F layers are necessarily discrete layers of ions, but rather that they may be gradients of ionization in a larger ionospheric scheme, with possibly a relatively homogeneous ionization between them which has a value of N (number of ions per cubic centimeter) probably not far from that obtaining at the top of

[24] Kenrick and Jen, Proc. I.R.E., vol. 17, pp. 718; April, (1929).
[25] Breit, Proc. I.R.E., vol. 17, p. 1509; September, (1929).
[26] Gilliland, Kenrick, and Norton, *Bur. Stan. Jour. Res.*, vol. 7, p. 1093; December, (1931).
[27] Epstein, *Proc. Nat. Acad. Sci.*, vol. 16, pp. 37 and 627; October, (1930).
[28] Appleton and Builder, *Proc. Phys. Soc.*, vol. 45, p. 208; March, (1933).

the E layer. The uncertainty as to the exact value and distribution of N in this "space" renders doubtful the meaning of values of retardation time found for the higher levels when translated into equivalent heights.

Most of the observations recent and current on ionospheric heights have converted the short base echo sounding method to a case where the transmitter and receiver are as close as in the same room, that is, the impulse travels vertically up to the layer and back. Then with a knowledge of equivalent height (computed from retardation time) and of the critical frequency, one sets n equal to zero[22,23,28] in Appleton's equation and solves for the value of N at that height. Since the critical frequency comes out differently for the ordinary and for the extraordinary rays, it turns out that the ordinary ray requires a value of N greater than that of the extraordinary ray to reduce the index of refraction to zero. Hence the ordinary ray will go higher before returning and will therefore be delayed behind the extraordinary ray in reaching the receiver.

In our case, for incidence upon the ionosphere at some angle, not $\pi/2$, where n does not have to vanish, it appears that the ordinary ray will still go higher than the extraordinary and there will be a difference in path length, and presumably therefore in phase, and a difference in angle of incidence at the receiver. The rays received as ordinary and extraordinary did not, of course, originate as the same ray, but one lost its ordinary component, which would come to earth beyond the receiver, while the other lost its extraordinary component, which would fall short of the receiver. While this phenomenon is likely to be a factor in causing the observed tilt angles, it is doubtful that the difference in angle of incidence would ever be great enough to cause them alone.

The last possibility which we shall examine is that of a wave which reaches the receiver after two deflections at the ionosphere and one at the ground, usually termed a "multiple." Such a ray would provide a $\Delta\theta$ of moderate size and, its path being not too much greater than that of the primary ray, would not need to be much more attenuated. It seems clear that when changes in the layer begin to take place this double contact ray might find it increasingly difficult to come through, and after the layer has risen might meet it at a very unfavorable angle.

Such multiples are quite usual on daytime records of layer heights, and Appleton and Barnett[29] found subsidiary maxima and minima on the interference patterns in their fringe system of observing, which they attributed to a ray twice deviated by the layer and once by the

[29] Appleton and Barnett, *Proc. Roy. Soc.*, ser. A, vol. 109, p. 621; December, (1925).

ground. It seems quite possible that some of the intermediate layers postulated by Colwell and Friend[11] and by Watson Watt *et al.*[12] may prove to be the result of multiples from the lower layers when account can be taken of the ionization gradients.

The clustering of angles, after the sunset period begins, in the region just below 180 degrees could, we have seen, be due either to the ground's becoming a better conductor or to a tendency toward a single wave at night. If we assume the former, it becomes very difficult to explain the sudden sallies into forbidden territory, but the latter explanation calls to mind the sudden appearance and disappearance of lower layers which one sees on night pulse records of layer heights. The point regarding the less favorable angle of incidence of the multiple upon a higher layer has been made in a previous paragraph.

If double refraction is the cause of the second wave by day, then the rising of the layer would have a result in the observed direction on two considerations. First, the higher layer would cause a smaller angle between the tops of the paths of ordinary and extraordinary rays as seen from the receiver, and second, the probable increase in the ionization gradient at the under side of the higher layer would tend to compress these tops toward each other.

Or again, if the second wave is due to a horizontally polarized ground wave, then the tremendous increase in sky wave around sunset could easily make its continued presence negligible.

Acknowledgment

It is a pleasure to the author to acknowledge not only the helpful counsel, but also the continued inspiration derived from Professor Ernest Merritt during the progress of this research.

Appreciation is also extended to the Heckscher Foundation for Research at Cornell University for the loan of apparatus used in the earlier part of the work and for the use of the tower as a field base.

Proceedings of the Institute of Radio Engineers
Volume 25, Number 11 *November, 1937*

TRANSMISSION THEORY OF PLANE ELECTROMAGNETIC WAVES*

BY
S. A. SCHELKUNOFF
(Bell Telephone Laboratories, Inc., New York City)

Summary—*This paper deals with transmission theory of plane electromagnetic waves in free space and in cylindrical regions of arbitrary cross section. Transmission properties of such waves can be expressed very simply in the same terms as the properties of electric waves guided by a pair of parallel wires. The earlier parts of the paper are concerned with general theorems and the latter parts with their application to plane waves in metal tubes of circular and rectangular cross sections.*

IN 1897, John William Strutt, Baron Rayleigh, published a paper, "On the Passage of Electric Waves Through Tubes, or Vibrations of Dielectric Cylinders" in which he showed that if the frequency is sufficiently high, electromagnetic waves can travel inside a conducting tube of either circular or rectangular cross section.[1] In 1898 R. C. MacLaurin obtained solutions appropriate to elliptic tubes. The papers that followed, duplicated to a large extent the results of Rayleigh and MacLaurin. All these early papers were concerned with wave propagation in perfectly conducting tubes; the attenuation of these waves was neglected.† The first papers dealing with attenuation and impedances appear to have been published only last year.[12,13,14] These papers were confined exclusively to cylindrical tubes of circular cross section.

In this paper the emphasis is placed on the general transmission theory of plane electromagnetic waves. The character of transmission in free space and in tubes of arbitrary cross section is fully described in terms of distributed series impedance and shunt admittance encountered by the waves. Then general formulas are derived for the attenuation constants due to dielectric losses as well as to absorption by the imperfectly conducting tubes. These general results are applied to tubes of circular and rectangular cross sections.

CLASSIFICATION OF ELECTROMAGNETIC WAVES

A word should be said concerning the precise meaning of the expression "plane wave." The more commonly accepted usage seems to be that the words *plane, spherical, cylindrical*, etc., refer to the shapes of equiphase surfaces, that is, surfaces of like phase at any given

* Decimal classification: R111. Original manuscript received by the Institute, February 25, 1937.
[1] Numbers refer to Bibliography.
† For a more complete historical background see reference (13).

instant. Occasionally these words are applied to electromagnetic fields represented by particular types of wave functions. We shall adhere to the more prevailing usage and call the waves plane if their equiphase surfaces form a family of parallel planes, spherical if the equiphase surfaces are concentric spheres, and cylindrical if the equiphase surfaces are coaxial cylinders. A ray is a straight line or a curve normal to the equiphase surfaces; it is the direction in which the phase of the wave varies most rapidly at any given instant. Thus the rays of a plane wave are parallel straight lines, those of a spherical wave the radii emanating from the center of the wave, and the rays of a cylindrical wave are the perpendicular radii emanating from the axis of the wave.

It has been customary to distinguish between longitudinal and transverse waves. In longitudinal waves the vibration takes place in the direction of the rays while in transverse waves it is at right angles to the rays. Contrary to the statements in most textbooks, electromagnetic waves are, in general, neither longitudinal nor transverse[16] in the sense that both vectors, E and H, are either longitudinal or transverse. There are some types of waves in which the electric vector is perpendicular to the ray; these waves will be called transverse electric waves. Similarly if the magnetic vector is perpendicular to the ray, then the waves will be called transverse magnetic waves.* Finally if both vectors are perpendicular to the ray, the wave is said to be transverse electromagnetic. Only the last of these conforms to the common conception of transverse waves. There are no electromagnetic waves in which either E or H is wholly in the direction of phase motion.

Transverse Magnetic Plane Waves in Free Space

Choose an equiphase surface as the xy plane. In accordance with the definition of transverse magnetic waves, the magnetic vector will be parallel to the xy plane so that H_z will vanish identically and the general electromagnetic equations will become†

$$\frac{\partial H_y}{\partial z} = -(g + i\omega\epsilon)E_x, \qquad \frac{\partial H_x}{\partial z} = (g + i\omega\epsilon)E_y,$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = (g + i\omega\epsilon)E_z, \tag{1}$$

* When dealing with waves traveling in cylindrical tubes, some investigators have used the names "H wave" and "E wave" instead of "transverse electric" and "transverse magnetic," respectively.

† The symbols g, μ, and ϵ designate respectively the conductivity, the permeability and the dielectric constant of the medium; $\omega = 2\pi f$, where f is the frequency. The field components are expressed as complex numbers and, in line with the existing usage, the time factor $e^{i\omega t}$ is dropped when unnecessary.

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -i\omega\mu H_x, \qquad \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -i\omega\mu H_y,$$

$$\frac{\partial E_y}{\partial x} = \frac{\partial E_x}{\partial y}, \qquad \frac{\partial H_x}{\partial x} + \frac{\partial H_y}{\partial y} = 0.$$

Since there is no magnetic flux in the direction of the rays, the electromotive force in any closed circuit contained entirely in an equiphase surface is zero. Hence the transverse component of the electromotive intensity can be regarded as the gradient of a potential function V.

$$E_t = -\operatorname{grad}_t V, \qquad E_x = -\frac{\partial V}{\partial x}, \qquad E_y = -\frac{\partial V}{\partial y}. \tag{2}$$

This function V represents the electromotive force acting between a given point and infinity along a path contained completely in the equiphase surface passing through the given point. We assume, of course, that the potential at infinity is zero; when such an assumption is inadmissible, V is the potential with respect to some other chosen reference point.

The last equation of (1) can be satisfied by setting

$$H_x = \frac{\partial A}{\partial y}, \qquad H_y = -\frac{\partial A}{\partial x}, \tag{3}$$

where A is of the nature of a stream function.* The flux of H across a curve PQ from right to left (Fig. 1) is

$$\begin{aligned}\Phi &= \int_{(PQ)} (H_y dx - H_x dy) = -\int_{(PQ)} \left(\frac{\partial A}{\partial x}\,dx + \frac{\partial A}{\partial y}\,dy\right) \\ &= A(P) - A(Q).\end{aligned} \tag{4}$$

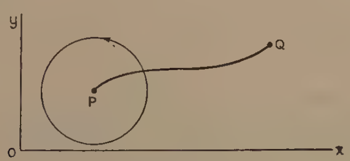


Fig. 1

Thus $A(P)$ represents the flux of H through a curve drawn from P to infinity or some other fixed reference point.

The potentials and the stream functions are not independent of each other; one relation follows immediately from the first two equa-

* If A is regarded as a vector parallel to the z axis, then $H = \text{curl } A$.

tions of (1) as soon as proper substitutions are made from (2) and (3)

$$\frac{\partial A}{\partial z} = -(g + i\omega\epsilon)V. \tag{5}$$

The longitudinal electric field can be obtained from (2) and from the third equation of (1)

$$E_z = -\frac{1}{g + i\omega\epsilon}\left(\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2}\right). \tag{6}$$

Another expression follows from the fourth and fifth equations of (1) if the values of the transverse field components are substituted in them; thus

$$E_z = -i\omega\mu A - \frac{\partial V}{\partial z} = \frac{1}{g + i\omega\epsilon}\left(\frac{\partial^2 A}{\partial z^2} - \sigma^2 A\right), \tag{7}$$

where $\sigma = \sqrt{i\omega\mu(g + i\omega\epsilon)}$. A comparison of (6) and (7) shows that A is a solution of the wave equation. Owing to (5), V is likewise a solution of the wave equation; thus

$$\Delta A = \frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} + \frac{\partial^2 A}{\partial z^2} = \sigma^2 A, \qquad \Delta V = \sigma^2 V. \tag{8}$$

We shall be interested in the particular solutions of the form $A = T(x, y)f(z)$ for which the transmission character in the direction of the z axis is the same for all rays. By our hypothesis the equiphase surfaces are parallel to the xy plane where $T(x, y)$ is real. This function T represents the relative distribution of the amplitude of the field over a typical equiphase surface. Substituting this in (8) and dividing by A, we obtain

$$\frac{1}{T}\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) + \frac{1}{f}\frac{d^2 f}{dz^2} = \sigma^2. \tag{9}$$

Inasmuch as the first term is independent of z and the remaining terms are independent of x and y, we must have

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = -\chi^2 T, \tag{10}$$

where χ is independent of x, y, z, and T. Since T is real, χ^2 must also be real.

Since A differs from T by a factor independent of x and y, A satisfies an equation of the same form as T. Then by (8)

$$\frac{\partial^2 A}{\partial z^2} - \sigma^2 A = -\frac{\partial^2 A}{\partial x^2} - \frac{\partial^2 A}{\partial y^2} = \chi^2 A \tag{11}$$

and the longitudinal electric field can be expressed as

$$E_z = \frac{\chi^2}{g + i\omega\epsilon} A. \tag{12}$$

Thus the longitudinal electric current density $(g+i\omega\epsilon)E_z$ differs from A only by a constant factor χ^2.

From (7) and (12) a second relationship between V and A is obtained

$$\frac{\partial V}{\partial z} = -\left(i\omega\mu + \frac{\chi^2}{g + i\omega\epsilon}\right)A. \tag{13}$$

Comparing (5) and (13) with the transmission equations of a smooth line

$$\frac{\partial V}{\partial z} = -ZI, \qquad \frac{\partial I}{\partial z} = -YV, \tag{14}$$

in which Z and Y are the distributed series impedance and shunt admittance of the line, we are led to call (5) and (13), the "transmission equations for transverse magnetic waves," and speak of

$$Z = i\omega\mu + \frac{\chi^2}{g + i\omega\epsilon}, \qquad Y = g + i\omega\epsilon, \tag{15}$$

as the series impedance and shunt admittance associated with waves of this type.* The analogy is all the more complete since V is the transverse electromotive force and A is proportional to the longitudinal electric current in a "wave tube" bounded by the rays.

The propagation constant and the characteristic impedance are

$$\Gamma_z = \sqrt{ZY} = \sqrt{\sigma^2 + \chi^2}, \qquad Z_z = \sqrt{\frac{Z}{Y}} = \frac{\Gamma_z}{g + i\omega\epsilon}. \tag{16}$$

Since g, μ, and ϵ are positive, Γ_z^2 is never below the real axis and the square root must be either in the first quadrant or in the third. We shall agree to designate by Γ that value of the square root which is in the first quadrant or on its boundaries and by $-\Gamma$ the remaining value (Fig. 2). The same convention applies to σ.

* A more general justification of the use of the impedance concept in this way will be found in a forthcoming paper by the author, "The impedance concept and its application to problems of reflection, refraction, shielding, and power absorption."

With the above convention in mind we write the following expressions for the progressive wave moving in the positive z direction

$$A = T(x, y)e^{-\Gamma_z z}, \qquad V = Z_z A, \tag{17}$$

in which $T(x, y)$ is the amplitude distribution in the plane $z=0$, the origin of time having been chosen to make the phase in this plane equal to zero. By (2), (3), and (12) the field components of this wave are

$$\begin{gathered} H_x = \frac{\partial T}{\partial y} e^{-\Gamma_z z}, \qquad H_y = -\frac{\partial T}{\partial x} e^{-\Gamma_z z}, \\ E_x = Z_z H_y, \qquad E_y = -Z_z H_x, \qquad E_z = \frac{\chi^2}{g + i\omega\epsilon} T e^{-\Gamma_z z}. \end{gathered} \tag{18}$$

For a wave moving in the opposite direction Γ_z and Z_z must be replaced by $-\Gamma_z$ and $-Z_z$, respectively.

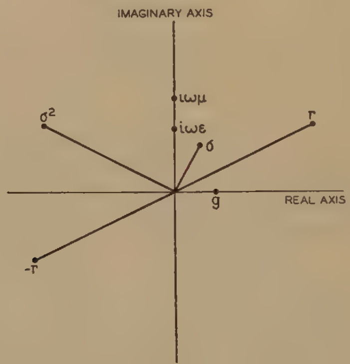


Fig. 2

Since T is real, H_x and H_y are in phase and their resultant has a definite direction at each point. Similarly, E_x and E_y are in phase and their resultant does not change its direction with time. Thus the magnetomotive intensity as well as the transverse component of the electromotive intensity is linearly polarized, although the direction of polarization varies from point to point. These two transverse field components are seen to be perpendicular to each other. The longitudinal component of the electromotive intensity is not necessarily in phase with the transverse component and the electric vector is, in general, elliptically polarized, with the plane of the ellipse normal to equiphase surfaces.

The equipotential lines in an equiphase surface are given by the following equation

$$T(x, y) = \text{constant}. \tag{19}$$

Since along these curves

$$\frac{\partial T}{\partial x}\,dx + \frac{\partial T}{\partial y}\,dy = 0, \tag{20}$$

their slope

$$\frac{dy}{dx} = -\frac{\dfrac{\partial T}{\partial x}}{\dfrac{\partial T}{\partial y}} \tag{21}$$

is the same as that of H. Hence (19) is also the equation of magnetic lines. The electric lines are three-dimensional curves and their projections on equiphase surfaces are orthogonal to magnetic lines.

Along a nodal line for T, E_z vanishes and so does the tangential component of the transverse electromotive intensity. Hence if the cylindrical surface formed by the lines parallel to the axis and passing through this nodal line is replaced by a perfectly conducting film, the boundary conditions will not be interfered with and we shall have plane waves in a bounded region. We shall return to this matter later.

We have seen that the relative amplitude distribution function T cannot be arbitrary but must satisfy a partial differential equation (10). The quantity χ depends upon T; in their turn, Γ_z and Z_z depend upon χ. Thus the propagation constant and the characteristic impedance of a transverse magnetic wave depend upon the electromagnetic properties of the medium, upon the frequency (that is upon the field distribution in time), and upon the character of amplitude distribution over equiphase surfaces. Comparing waves with different amplitude distribution but corresponding to the same χ, we find them traveling with the same velocity, having the same attenuation and the same characteristic impedance; provided of course that the frequency and the medium are the same. For all practical purposes they will merge into one wave. On the other hand, if the χ's of the waves are different, the waves will either "run away" from each other or one will subside faster than the other. Thus we come to look upon T and χ as essential characteristics of a plane wave. Later we shall find that for plane waves traveling in a metal tube, T and χ depend upon the shape and the size of the cross section of the tube.

We should add, perhaps, that over any single equiphase surface, the amplitude distribution may be perfectly arbitrary; there is nothing to prevent us from enforcing it by a proper distribution of applied forces. But this relative amplitude distribution will not be preserved as the wave proceeds on its way (without the aid, of course, of forces

applied all along the wave path). What happens is that the original amplitude distribution may be regarded as the resultant of several or an infinite number of the "permissible" distributions satisfying (10), the latter corresponding to different χ's and hence proceeding on their way with different attenuation and different velocities. The analysis of a given spatial amplitude distribution into canonical distributions satisfying (10) is analogous to Fourier analysis of an arbitrary time function and has the same purpose. Such important ideas as the impedance and the propagation constant have a meaning only for certain types of temporal and spatial distributions of forces. These particular distributions are frequently called harmonics, more specifically time harmonics and space harmonics.

We now shall concentrate our attention on a wave with a given amplitude distribution $T(x, y)$ and inquire what happens to the propagation constant and the characteristic impedance as the frequency is varied. Let us start with an important class of waves for which $\chi = 0$. In this case the longitudinal electric field vanishes by (12) and the waves are transverse electromagnetic. The function T satisfies Laplace's equation in two dimensions so that electric and magnetic lines are exactly the same as in the corresponding two-dimensional static cases. If any electrostatic pattern is forced to vary its intensity with some frequency f, this pattern will move with a certain velocity at right angles to itself. Ordinary plane waves of light are of this type; waves guided by a pair of perfectly conducting wires or a pair of coaxial cylinders are also of this type. For waves guided by imperfectly conducting wires an allowance has to be made for additional dissipation caused by the wires draining the energy out of the waves. In this case a small longitudinal field makes it appearance, thereby making the waves only approximately transverse electromagnetic.

The propagation constant of transverse electromagnetic waves is σ; their characteristic impedance will be designated by η. These constants are so closely related to the fundamental constants g, μ, and ϵ and so frequently appear as entities in equations for all kinds of electromagnetic waves that they may themselves be regarded as fundamental or intrinsic constants of the medium. The intrinsic propagation constant and the intrinsic impedance of the medium are defined, therefore, by

$$\sigma = \alpha + i\beta = \sqrt{i\omega\mu(g + i\omega\epsilon)},$$
$$\eta = \sqrt{\frac{i\omega\mu}{g + i\omega\epsilon}} = \frac{\sigma}{g + i\omega\epsilon} = \frac{i\omega\mu}{\sigma}. \tag{22}$$

The real part of σ is the attenuation constant and β is the phase constant.

Since the field along a ray varies as $e^{-\sigma z+i\omega t}=e^{-\alpha z}e^{i(\omega t-\beta z)}$, the phase of the wave appears constant to an observer moving along the ray with a velocity $c=\omega/\beta$ as obtained from the condition

$$d(\omega t-\beta z)=\omega dt-\beta dz=0, \qquad c=\frac{dz}{dt}=\frac{\omega}{\beta}. \tag{23}$$

This is the so-called phase velocity of the wave or simply the wave velocity.

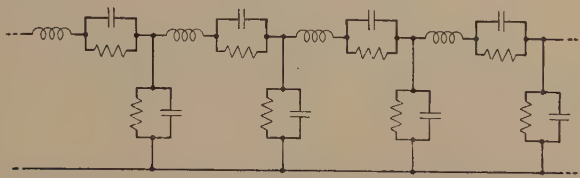


Fig. 3—A representation of the transmission characteristics of transverse magnetic waves. All the circuit parameters are supposed to be continuously distributed.

The wave length is defined as the distance λ from crest to crest. Since the phase change from crest to crest is 2π, we have

$$\beta\lambda=2\pi \quad \text{and} \quad \lambda=\frac{2\pi}{\beta} \quad \text{or} \quad \beta=\frac{2\pi}{\lambda}. \tag{24}$$

From (23) and (24) we obtain the familiar relation between the wave velocity, the frequency, and the wave length

$$f\lambda=c. \tag{25}$$

If the medium is nondissipative, $g=0$ and

$$\sigma=i\omega\sqrt{\mu\epsilon}, \quad \beta=\omega\sqrt{\mu\epsilon}, \quad c=\frac{1}{\sqrt{\mu\epsilon}}, \quad \eta=\sqrt{\frac{\mu}{\epsilon}}, \tag{26}$$

so that the wave velocity and the impedance of transverse electromagnetic waves are independent of the frequency.

Passing to the general case of $\chi\neq0$, we recall that for plane waves χ^2 must be real and consequently χ is either real or a pure imaginary. In the former case χ^2 is positive and the distributed impedances Z and Y can be regarded as a combination of inductances, capacitances, and resistances as shown in Fig. 3. Such an equivalent network is a high-pass filter.*

* If the elements were lumped, the structure would have been a band-pass filter; but the upper cutoff recedes to infinity as the structure becomes more fine-grained.

For the present we shall assume that χ is real and that the medium is nondissipative. The propagation constant and the characteristic impedance (16) become

$$\Gamma_z = \sqrt{\chi^2 - \omega^2\mu\epsilon}, \qquad Z_z = \frac{\sqrt{\chi^2 - \omega^2\mu\epsilon}}{i\omega\epsilon}. \tag{27}$$

It is now evident that when ω is sufficiently small, Γ_z is real and Z_z is a pure imaginary. Consequently the wave is exponentially attenuated and on the average there is no flow of energy across any particular

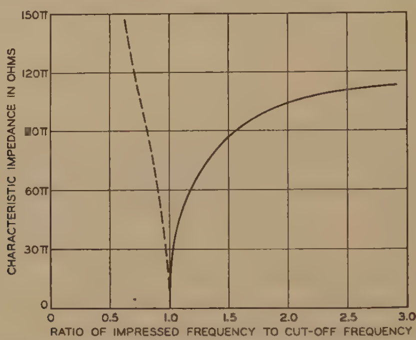


Fig. 4—The characteristic impedance of transverse magnetic waves in air. The solid curve represents the resistance component and the dotted curve the negative of the reactance component.

equiphase surface. At sufficiently high frequencies Γ_z is a pure imaginary and Z_z is real; then there exists a steady, on the average, flow of energy in the direction of the wave.

The critical or cutoff frequency is the frequency at which the propagation constant vanishes. This frequency f_c and the corresponding wave length are

$$\omega_c = \frac{\chi}{\sqrt{\mu\epsilon}}, \qquad f_c = \frac{\chi}{2\pi\sqrt{\mu\epsilon}}, \qquad \lambda_c = \frac{2\pi}{\chi}. \tag{28}$$

Introducing ω_c in place of χ into (27), we have

$$\begin{aligned} \Gamma_z &= \sqrt{(\omega_c^2 - \omega^2)\mu\epsilon}, \quad Z_z = -i\sqrt{\left(\frac{\omega_c^2}{\omega^2} - 1\right)\frac{\mu}{\epsilon}}, \text{ if } \omega < \omega_c, \\ \Gamma_z &= i\sqrt{(\omega^2 - \omega_c^2)\mu\epsilon}, \quad Z_z = \sqrt{\left(1 - \frac{\omega_c^2}{\omega^2}\right)\frac{\mu}{\epsilon}}, \text{ if } \omega > \omega_c. \end{aligned} \tag{29}$$

In Fig. 4 are shown the impedance characteristics for transverse magnetic waves in air.

The cutoff frequency and the cutoff wave length depend upon χ and therefore upon the character of amplitude distribution. For example, if the distribution of the amplitude over equiphase surfaces is given by $T(x, y) = \cos 5x/a$ or $T(x, y) = 3 \sin (3x/a) \cos (4y/a)$ where a is some length, $\chi = 5/a$ and $\lambda_c = 2\pi a/5$. For transverse electromagnetic waves χ vanishes and consequently $\lambda_c = \infty$, $f_c = 0$.

From (18) we find that in nondissipative media E_z is in phase with the transverse electric field below the cutoff and in quadrature above the cutoff. Hence, the electric field is elliptically polarized above the cutoff and linearly polarized below it.

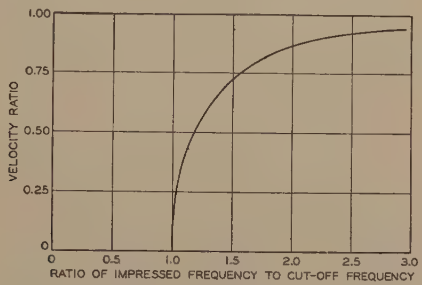


Fig. 5—The ratio of the velocity of light to the velocity of transverse magnetic waves.

Let the ratio of the cutoff frequency to the applied frequency be

$$\nu = \frac{f_c}{f} = \frac{\omega_c}{\omega} = \frac{\lambda}{\lambda_c}. \tag{30}$$

Introducing this into (29), we have above the cutoff

$$\Gamma_z = i\beta\sqrt{1 - \nu^2}, \qquad Z_z = \eta(1 - \nu^2)^{1/2}. \tag{31}$$

The phase velocity c_z is the velocity with which an observer must advance along a ray in order to keep abreast of the phase and the wave length λ_z is the distance from crest to crest; thus

$$\Gamma_z = i\beta_z = \frac{i\omega}{c_z}, \qquad \lambda_z f = c_z. \tag{32}$$

Comparing (31) and (32), we obtain

$$\beta_z = \beta\sqrt{1 - \nu^2}, \qquad c_z = \frac{c}{\sqrt{1 - \nu^2}}, \qquad \lambda_z = \frac{\lambda}{\sqrt{1 - \nu^2}}. \tag{33}$$

The ratio of the phase velocity of transverse electromagnetic waves to that of transverse magnetic waves (with real χ) is shown in Fig. 5.

The group velocity is defined by

$$v = \frac{d\omega}{d\beta_z}, \tag{34}$$

where β_z is the phase constant of the wave. From the preceding equation we obtain the group velocity

$$v = \frac{\beta_z}{\omega\mu\epsilon} = \frac{c^2}{c_z} = c\sqrt{1 - \nu^2}. \tag{35}$$

If χ is a pure imaginary, χ^2 is negative and there is no cutoff and the wave transmits energy at all frequencies. Later we shall find that for plane waves in a metal tube χ must be real and for this reason the case of imaginary χ is of lesser interest to us.

Returning to the dissipative case, we find that strictly speaking there is no cutoff and, in a way, all frequencies are transmitted. At low frequencies, however, the amount of transmitted power is small. Not until the frequency begins to exceed the cutoff determined by disregarding the losses, does the amount of transmitted power become appreciable. It is natural, therefore, to speak of the cutoff in the dissipative case as well and define it by (28). Above the cutoff we have

$$\begin{aligned} \Gamma &= \sqrt{\sigma^2 + \chi^2} = \sqrt{\omega_c^2\mu\epsilon - \omega^2\mu\epsilon + i\omega\mu g} \\ &= i\omega\sqrt{\mu\epsilon}\sqrt{(1 - \nu^2) - i\frac{g}{\omega\epsilon}}\,. \end{aligned} \tag{36}$$

If the "Q" of the dielectric, defined as the ratio $\omega\epsilon/g$, is large, then sufficiently above the cutoff we have approximately

$$\Gamma_z = \alpha_z + i\beta_z = \tfrac{1}{2}g\sqrt{\frac{\mu}{\epsilon}}(1 - \nu^2)^{-1/2} + i\omega\sqrt{\mu\epsilon(1 - \nu^2)}. \tag{37}$$

Ultimately, at frequencies very high above the cutoff, the attenuation constant is

$$\alpha_z = \tfrac{1}{2}g\sqrt{\frac{\mu}{\epsilon}}\,. \tag{38}$$

The exact formulas for the attenuation and the phase constants are

$$\begin{aligned} \alpha_z &= \sqrt{\frac{(\omega_c^2 - \omega^2)\mu\epsilon + \sqrt{(\omega_c^2 - \omega^2)^2\mu^2\epsilon^2 + \omega^2\mu^2 g^2}}{2}}\,, \\ \beta_z &= \sqrt{\frac{(\omega^2 - \omega_c^2)\mu\epsilon + \sqrt{(\omega_c^2 - \omega^2)^2\mu^2\epsilon^2 + \omega^2\mu^2 g^2}}{2}}\,. \end{aligned} \tag{39}$$

Returning to (16), we observe that for sufficiently high frequencies we have approximately

$$\Gamma_z = \sigma\left(1 + \frac{\chi^2}{2\sigma^2}\right) = \sigma + \frac{\chi^2}{2\sigma} \cdot \tag{40}$$

Ultimately as the frequency increases, Γ_z approaches the intrinsic propagation constant of the medium and thus it becomes substantially independent of the character of the amplitude distribution over equiphase surfaces.

Likewise, the characteristic impedance and hence the ratio of the transverse field components approach a definite limit with increasing frequency. The longitudinal component, on the other hand, diminishes with increasing frequency and the wave tends to become transverse electromagnetic.

Transverse Magnetic Plane Waves and Ordinary Plane Waves of Light

The most familiar "plane wave" is the wave in which E (and H as well) has the same amplitude over a typical equiphase plane and which on this account can properly be called uniform. If, at the same time, the directions of E and H are independent of time, then the wave is a linearly polarized* uniform plane wave; in what follows such a wave will be designated simply as an "ordinary wave," the complete description being too long.

It is easily seen that several ordinary waves can be superposed in such a way as to produce a transverse magnetic wave. Consider a wave of unit amplitude, traveling with velocity c in the direction making angles A, B, and C with the co-ordinate axes (Fig. 6). The electric field of such a wave is represented by

$$E = \exp i\omega\left(t - \frac{x \cos A + y \cos B + z \cos C}{c}\right). \tag{41}$$

The expression $x \cos A + y \cos B + z \cos C$ is simply the distance of a typical point from the plane $x \cos A + y \cos B + z \cos C = 0$. Let H be parallel to the xy plane; then E makes an angle (90 degrees $- C$) with the z axis. Taking another wave of unit amplitude, traveling in the direction making angles 180 degrees $- A$, 180 degrees $- B$ and C with the co-ordinate axis and adding it to the first, we obtain

* Among the physicists these waves are known as "plane polarized." The engineering usage here adopted seems to us more appropriate.

$$E_z = \sin C\left[\exp i\omega\left(t - \frac{x\cos A + y\cos B}{c} - \frac{z\cos C}{c}\right) + \exp i\omega\left(t + \frac{x\cos A + y\cos B}{c} - \frac{z\cos C}{c}\right)\right]$$
$$= 2\sin C\cos\frac{x\cos A + y\cos B}{c}\exp i\omega\left(t - \frac{z\cos C}{c}\right). \tag{42}$$

Only the amplitude of this wave varies with x and y while the phase

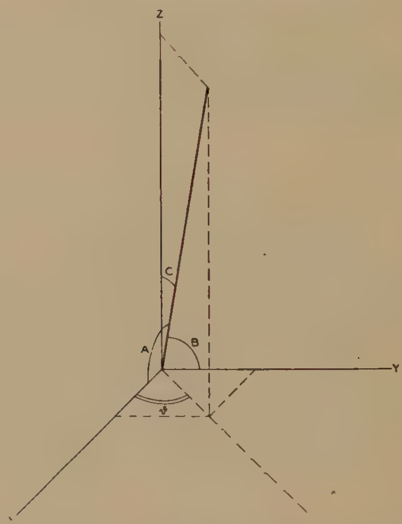


Fig. 6

depends only upon z and t. This amplitude pattern travels in the z direction with the velocity $c/\cos A$.

In "synthesizing" more complicated transverse magnetic waves out of ordinary plane waves, we see to it that the wave directions lie on a cone (Fig. 7) whose generators make a constant angle C with a fixed line (the z axis, for instance) and that the magnetic vectors are parallel to a fixed plane (the xy plane, for example). In the case of wave directions making different angles with the z axis, the phases of the component waves as observed along the z axis would progress at different rates and the resultant could not be regarded as a single wave advancing along the axis with a definite velocity. Particular wave directions may be specified by a single co-ordinate θ, θ being the inclination of the projection of the wave direction upon the xy plane (Fig. 6). Evidently

$$\cos A = \sin C\cos\theta,\ \cos B = \sin C\sin\theta. \tag{43}$$

The two components of (42) correspond to θ and 180 degrees$-\theta$, that is, to two diametrically opposite wave directions on the cone.

When an infinite number of elementary waves are added together, their amplitudes must be taken indefinitely small. Let the amplitude of the element traveling in the direction specified by the parameter θ be $A(\theta)d\theta$. The longitudinal electric field of the resultant wave is

$$\begin{aligned} E_z &= \sin C \int_0^{2\pi} A(\theta) \exp i\omega \left[t - \frac{(x \cos \theta + y \sin \theta) \sin C + z \cos C}{c} \right] d\theta \\ &= \sin C \exp i\omega \left(t - \frac{z \cos C}{c} \right) \times \\ &\qquad \int_0^{2\pi} A(\theta) \exp \frac{-i\omega(x \cos \theta + y \sin \theta) \sin C}{c} d\theta. \end{aligned} \tag{44}$$

In relation to the preceding section, this wave is characterized by $\chi = (\omega/c) \sin C = \beta \sin C$ and $\Gamma_z = (\omega/c) \cos C = \beta \cos C$.

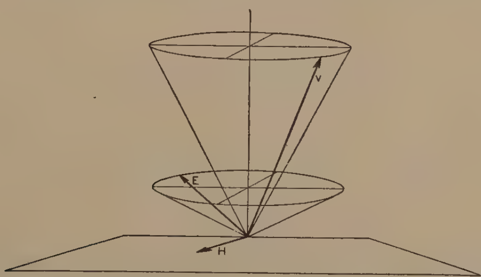


Fig. 7—The cone of wave directions for a bundle of ordinary plane waves with their magnetic vectors parallel to a fixed plane.

If this wave is to possess circular symmetry, we must choose all elements equal and set $A(\theta) = 1$. Indeed, under these conditions it can be shown that

$$E_z = 2\pi \sin C \, J_0\left(\frac{\omega\rho \sin C}{c}\right) \exp i\omega \left(t - \frac{z \cos C}{c} \right), \tag{45}$$

where J_0 is the Bessel function of order zero. Similarly if the amplitudes of the elements are equal but the phases are given by $e^{in\theta}$, where n is an integer, then

$$E_z = 2\pi \sin C \, J_n\left(\frac{\omega\rho \sin C}{c}\right) \cos n\phi \exp i\omega \left(t - \frac{z \cos C}{c} \right). \tag{46}$$

The surfaces along which E_z vanishes are concentric cylinders and equi-spaced radial planes. It is for this reason that portions of such waves can be "squeezed" into circular metal tubes.

Not all transverse magnetic waves can be regarded as finite or infinite bundles of ordinary plane waves. This is evident from the fact that under no circumstances can (44) possibly represent a wave attenuated exponentially in the z direction. A closely related observation was made by T. C. Fry with regard to impossibility of decomposing "hybrid light" into ordinary light waves.[16] The most serious limitation is, however, that in dissipative media transverse magnetic waves can never be decomposed into ordinary plane waves.* Usually, however, these waves can be represented as bundles of "exponential plane waves" considered by Fry.

Generation of Transverse Magnetic Waves

Strictly speaking no plane wave in unlimited space can be generated by a source which is finite in extent and, therefore, no plane wave can be generated under practical conditions. On the other hand, it is usually pos-

Fig. 8

sible to approximate plane waves in sufficiently limited regions. Thus, at great distances from an antenna radiating in free space, the wave is substantially an ordinarily plane wave provided our observations are limited to a sufficiently small volume. We must limit the observed area of the equiphase surface or else the nonuniformity of the field will become noticeable and we must limit the radial depth of the observed region or else the spherical divergence of the wave may be appreciable.

From the considerations of the preceding section, it is evident that some transverse magnetic waves can also be approximated by a proper arrangement of antennas. Consider two widely separated parallel radi-

* This insufficiency of optical "explanation" of transverse magnetic waves has apparently escaped L. Brillouin's[15] attention. Thus, on p. 239 of his paper he makes the following strong statement: "Our investigation has shown clearly the precise nature of waves which can be propagated in a tube; they are systems of interference bands which are produced during the reflection of a plane wave in a system of mirrors. All the very special characteristics of these waves follow from this statement." The author wishes to point out that even when transverse magnetic waves can be regarded as bundles of ordinary plane waves, the elementary components *are not necessarily* the reflected waves of some one component in a system of mirrors.

ating wires A and B (Fig. 8). Let the sources be in phase and let their amplitudes be equal. In a sufficiently limited stretch along the perpendicular OC, the two waves from A and B are substantially ordinary plane waves and their resultant is approximately a transverse magnetic plane wave. The phase velocity is very large in regions near O and it diminishes gradually as we recede from the sources.

Several sources equidistributed on a circle of radius large compared with the wave length may be used to produce transverse magnetic waves with nearly circular symmetry. Calculations show that even six sources make a good approximation sufficiently near the axis of the circle.

In the presence of a perfectly reflecting plane, electromagnetic waves are of the ordinary plane wave type only at the so-called grazing incidence. Otherwise the reflected wave combines with the "direct" wave to form a new wave that is not completely transverse. If the magnetic vector of the incident wave is parallel to the reflector, the resultant wave is transverse magnetic. While we habitually speak of a "direct" and a "reflected" wave, we should remember, of course, that under the present conditions one cannot exist without the other and that consequently their resultant is more "real" than the components.* The usual method of thought about reflection is analogous to thinking about 5 as 2+3. That this habit of thought may be very convenient on some occasions, is not to be disputed; nevertheless, this manner of thought represents only one aspect of the reality.

It has been pointed out that exponentially attentuated transverse magnetic waves cannot be regarded as interference patterns of ordinary plane waves. Such exponentially attenuated waves can be generated, however, by means of ordinary plane waves shining on the interface between two semi-infinite media, one of which is nondissipative and has lower intrinsic velocity (larger $\mu\epsilon$) than the other. A bundle of waves (44) can be utilized to produce a certain amplitude pattern $T(x, y)$ in the medium with the lower intrinsic velocity. This pattern is preserved by the reflection from the interface and thus it is impressed on the other medium. The properties of the other medium may be such that this pattern will have no alternative but to be attenuated exponentially with the distance from the interface. To prove that this can happen we merely have to refer to the phenomenon of total internal reflection.

A perfectly general theoretical way of generating transverse magnetic waves with any given amplitude pattern is by means of plane electric current sheets. Across an electric current sheet the tangential

* We are thinking, here, in terms of the steady-state conditions.

component of E is continuous and the discontinuity in the tangential component of H is equal to the electric current density. If the sheet is in a homogeneous medium, the tangential components of H on the two sides of the sheet must be equal but oppositely directed.

For instance, the wave given by (18) can be produced by means of a current sheet in the xy plane. In accordance with Ampère's law the current densities of the sheet must be

$$J_x = 2\frac{\partial T}{\partial x}, \qquad J_y = 2\frac{\partial T}{\partial y}. \tag{47}$$

Below the sheet there will be a wave similar to (18) but moving in the opposite direction. This wave can be removed by a reflector properly placed below the current sheet.

Transverse Electric Plane Waves

The general theory of transverse electric waves is very similar to that of transverse magnetic waves. The electric and magnetic vectors simply change their rôles. Thus, the general electromagnetic equations become

$$\begin{gathered}
\frac{\partial E_y}{\partial z} = i\omega\mu H_x, \qquad \frac{\partial E_x}{\partial z} = -i\omega\mu H_y, \qquad E_z = 0, \\
\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -i\omega\mu H_z, \\
\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = (g + i\omega\epsilon)E_x, \qquad \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = (g + i\omega\epsilon)E_y, \\
\frac{\partial H_y}{\partial x} = \frac{\partial H_x}{\partial y}, \qquad \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = 0.
\end{gathered} \tag{48}$$

As there is no electric current in the direction of the rays, the magnetomotive force in any closed circuit contained entirely in an equiphase surface is zero. Hence, the transverse component of the magnetomotive intensity can be regarded as the gradient of a magnetic potential function U

$$H_t = -\operatorname{grad}_t U, \qquad H_x = -\frac{\partial U}{\partial x}, \qquad H_y = -\frac{\partial U}{\partial y}. \tag{49}$$

This function U represents the magnetomotive force between a given point and infinity (or some other fixed point) along a path contained completely in the equiphase surface passing through the given point.

The electromotive intensity is derivable from a stream function F

$$E_x = -\frac{\partial F}{\partial y}, \qquad E_y = \frac{\partial F}{\partial x}. \tag{50}$$

If this function is regarded as a vector potential parallel to the z axis, then $E = -\text{curl } F$. The electric flux across PQ (Fig. 1) from left to right is

$$\Phi = \int_{(PQ)} (E_x dy - E_y dx) = F(P) - F(Q) \tag{51}$$

so that F may be interpreted as the electric flux through a curve drawn from a given point to infinity or some other fixed point. This curve must lie, of course, in the equiphase surface passing through the given point.

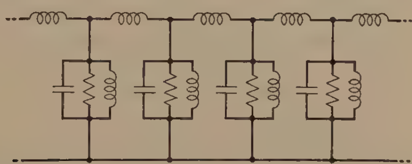


Fig. 9—A representation of the transmission characteristics of transverse electric waves. All the circuit parameters are supposed to be continuously distributed.

Following the familiar line of argument, we shall arrive at the following equations

$$\frac{\partial F}{\partial z} = -ZU, \qquad \frac{\partial U}{\partial z} = -YF, \tag{52}$$

where,

$$Z = i\omega\mu, \qquad Y = (g + i\omega\epsilon) + \frac{\chi^2}{i\omega\mu}. \tag{53}$$

Equations (52) are similar to the equations connecting the voltage and the current in a transmission line having Z for its distributed series impedance and Y for the shunt admittance. For this reason we may call (52) the transmission equations of transverse electric waves. The nature of this transmission is represented in Fig. 9.

The functions U and F are proportional to $T(x, y)$, governing amplitude distribution over equiphase surfaces. This function satisfies (10). The complete expressions for the field of a progressive transverse electric plane wave are as follows

$$U = T(x, y)e^{-\Gamma_z z}, \qquad F = Z_z U, \qquad Z_z = \frac{i\omega\mu}{\Gamma_z},$$

$$H_x = -\frac{\partial T}{\partial x} e^{-\Gamma_z z}, \qquad H_y = -\frac{\partial T}{\partial y} e^{-\Gamma_z z}, \tag{54}$$

$$E_x = Z_z H_y, \qquad E_y = -Z_z H_x,$$

$$H_z = \frac{\chi^2}{i\omega\mu} F = \frac{\chi^2}{\Gamma_z} T e^{-\Gamma_z z}.$$

Inasmuch as the propagation constant is given by exactly the same expression as in the case of transverse magnetic waves, the equations beginning with (27) and ending with (39), excepting those for the impedance, are equally applicable to transverse electric waves. But the characteristic impedance behaves differently; it becomes infinite rather than zero at the cutoff and otherwise it varies as

$$Z_z = \sqrt{\frac{\mu}{\epsilon}}\,(1 - \nu^2)^{-1/2}. \tag{55}$$

From (54) we conclude that H_x and H_y are always in phase and that consequently they can be combined into a single vector whose direction is fixed in time but may vary from point to point. The same is true with regard to E_x and E_y. Thus, in transverse electric waves, the electric vector and the transverse component of the magnetic vector are linearly polarized; these two vectors are mutually perpendicular. In nondissipative media, the longitudinal component of H is in phase with the remaining components below the cutoff where Γ_z is real and in quadrature above the cutoff where Γ_z is a pure imaginary. Hence, in nondissipative media the complete magnetic vector is linearly polarized below the cutoff and elliptically polarized above the cutoff. In dissipative media the magnetic vector is always elliptically polarized.

As the frequency increases indefinitely, Γ_z also increases indefinitely and consequently H_z diminishes indefinitely. Hence, transverse electric waves tend to become transverse electromagnetic if the frequency is allowed to increase without limit.

For transverse electric waves the curves given by (19) represent simultaneously magnetic equipotential lines and electric lines. If a perfectly conducting cylindrical film is placed normally to such lines, the boundary conditions are not interfered with and we shall have transverse electric waves in a bounded region.

A simple relationship exists between the characteristic impedances of transverse magnetic and transverse electric waves; this is

$$Z_{TM}Z_{TE} = \eta^2, \qquad \eta = \sqrt{\frac{i\omega\mu}{g + i\omega\epsilon}}. \tag{56}$$

The product of the characteristic impedances of transverse electric and transverse magnetic waves corresponding to the same value of χ is equal to the square of the intrinsic impedance of the medium.

A bundle of uniform linearly polarized plane waves traveling in directions lying on a circular cone, so oriented that their electric vectors are parallel to a fixed plane, may combine into a transverse electric wave. If the xy plane is chosen as the fixed plane, H_z will be given by the expression on the right of (44). Since the integral is proportional to $T(x, y)$, the amplitudes $A(\theta)d\theta$ of the elementary waves must be such that the integral is either real or differs from a real function by a constant factor; otherwise the resultant wave will not be plane.

Plane Waves in Cylindrical Tubes

It has been already intimated that plane waves can exist inside perfectly conducting cylindrical tubes. In metallic tubes the waves will be only approximately plane, the equiphase surfaces being distorted at the boundaries. Inasmuch as we are interested only in transmission of waves through well-conducting tubes, we shall assume them to be perfect to begin with and then calculate the principal correction in the form of the attentuation constant.

The boundary conditions are then the vanishing of the tangential component of E or the normal component of H. In transverse magnetic waves the longitudinal component of E is proportional to the amplitude distribution function T. Thus, the boundary condition is $T=0$ at the surface of the tube. Since the transverse electric field is proportional to the gradient of T, the tangential component of the transverse field vanishes automatically when $T=0$. For transverse electric waves the normal component of H is proportional to the normal derivative $\partial T/\partial n$ of the amplitude distribution function and the boundary condition becomes $\partial T/\partial n=0$ at the surface of the tube.

The equation satisfied by T is seen to be identical with that for the amplitude of an elastic membrane oscillating in one of its natural modes. This simple observation is valuable in visualizing the stable patterns of electric field distribution. There is no reason why we could not assume arbitrary distribution of the transverse field over some particular cross section of the tube but this particular pattern will not, in general, move along the tube without changing shape. Only those

patterns will preserve themselves in motion through the tube for which T is a solution of (10). And these patterns are the patterns assumed by the displacement of an elastic membrane, of the same shape as the cross section of the tube, oscillating in one of its natural modes. It may be observed that the longitudinal electric current density in transverse magnetic waves and the longitudinal magnetic current density in transverse electric waves are proportional to T, which makes the analogy quite close. At the cutoff either E or H is wholly longitudinal and we have simply a free oscillation of electromagnetic energy in the tube, transversely to it. As the frequency is made to exceed this critical frequency and the same pattern of amplitude distribution is enforced across some cross section, the field is compelled to move to preserve its electrodynamic equilibrium. Transverse magnetic waves correspond to oscillations in a membrane clamped around the edge while transverse electric waves correspond to oscillations of a membrane with a free edge.

It has been shown that for plane waves the constant χ^2 is real. We shall now prove that for waves inside a perfectly conducting tube not only χ^2 but χ itself is real and this is regardless of the particular shape of equiphase surfaces. The truth of this theorem follows at once from Green's theorem

$$\iint_{(S)} \left(\frac{\partial U}{\partial x}\frac{\partial V}{\partial x} + \frac{\partial U}{\partial y}\frac{\partial V}{\partial y}\right) dS = -\int_{(s)} U \frac{\partial V}{\partial n} ds - \iint_{(S)} U\left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2}\right) dS, \tag{57}$$

where the double integration is extended over the area S and the simple integration over its periphery. The two functions U and V are assumed to be continuous and twice differentiable over S. The normal is drawn outwards from the area S. Let T^* be the conjugate complex† of T and apply Green's theorem and (10) to obtain

$$\iint_{(S)} \left(\frac{\partial T}{\partial x}\frac{\partial T^*}{\partial x} + \frac{\partial T}{\partial y}\frac{\partial T^*}{\partial y}\right) dxdy = -\int_{(s)} T^* \frac{\partial T}{\partial n} ds + \chi^2 \iint_{(S)} TT^* dxdy. \tag{58}$$

At the surface of the tube either T or $\partial T/\partial n$ vanishes so that the first term on the right disappears. The two remaining integrals are essen-

† Since in this discussion we do not assume that the waves are plane, $T(x, y)$ may be complex.

tially positive; therefore, χ^2 is positive and χ must be real. There is no loss in generality in assuming that χ is positive.

The above conclusion is based essentially upon the existence of closed nodal lines and hence it is applicable to free-space waves possessing this property. Consequently, if χ is a pure imaginary, the nodal lines are either open or nonexistent. Of course, there is nothing in the argument to prevent χ from being real even if nodal lines are open.

When the waves are plane, T itself must be real and $T^* = T$; consequently

$$\chi^2 = \frac{\iint_{(S)} |\operatorname{grad} T|^2 dS}{\iint_{(S)} T^2 dS}. \tag{59}$$

The constant χ depends upon the "degree of smoothness" of the amplitude pattern, that is, upon the relative gradient $|\operatorname{grad} T|/T$. If the latter is everywhere small, χ is small and the cutoff frequency is low.

The following similarity transformation

$$u = x\sqrt{\frac{\pi}{S}}, \qquad v = y\sqrt{\frac{\pi}{S}}, \tag{60}$$

changes the area S of the cross section and makes it equal to the area of a circle of unit radius. Applying this transformation to (59), we obtain

$$\chi = k\sqrt{\frac{\pi}{S}}, \tag{61}$$

where,

$$k^2 = \frac{\iint \left[\left(\frac{\partial T}{\partial u}\right)^2 + \left(\frac{\partial T}{\partial v}\right)^2\right] du dv}{\iint T^2 du dv}. \tag{62}$$

The constant k is independent of the size of the tube and is a function of its shape and the particular transmission mode, that is, of the particular amplitude pattern. For this reason k will be called the modular constant.

We now determine the average magnetic energy W^m stored per unit length of the tube in a transverse magnetic wave, the average energy W_t^e associated with the transverse electric field and the average energy W_l^e associated with the longitudinal electric field

$$W^m = \tfrac{1}{4}\mu \iint_{(S)} HH^* dS = \tfrac{1}{4}\mu \iint_{(S)} |\operatorname{grad} T|^2 dS$$

$$W_t^e = \tfrac{1}{4}\epsilon \iint_{(S)} E_t E_t^* dS = \tfrac{1}{4}\mu(1 - \nu^2) \iint_{(S)} |\operatorname{grad} T|^2 dS$$

$$W_l^e = \tfrac{1}{4}\epsilon \iint_{(S)} E_z E_z^* dS = \tfrac{1}{4}\mu\nu^2\chi^2 \iint_{(S)} T^2 dS. \tag{63}$$

By virtue of (59) the last equation becomes

$$W_l^e = \tfrac{1}{4}\mu\nu^2 \iint_{(S)} |\operatorname{grad} T|^2 dS. \tag{64}$$

Thus the average magnetic energy in a transverse magnetic wave is equal to the average electric energy

$$W^m = W_t^e + W_l^e. \tag{65}$$

At the cutoff the entire electric energy is associated with the longitudinal electric field and sufficiently above the cutoff most of it is associated with the transverse field.

Similarly, we have for the average energy associated with different field components of a transverse electric wave

$$W_t^m = \tfrac{1}{4}\mu(1 - \nu^2) \iint_{(S)} |\operatorname{grad} T|^2 dS,$$

$$W_l^m = \tfrac{1}{4}\mu\nu^2 \iint_{(S)} |\operatorname{grad} T|^2 dS, \tag{66}$$

$$W^e = \tfrac{1}{4}\mu \iint_{(S)} |\operatorname{grad} T|^2 dS.$$

Here also the average magnetic and electric energies are equal.

$$W_t^m + W_l^m = W^e. \tag{67}$$

At the cutoff the magnetic energy is associated exclusively with the longitudinal magnetic field and sufficiently high above the cutoff most of the energy is in the transverse field.

The average power flow in the direction of the tube is by the complex Poynting flux theorem

$$W = \tfrac{1}{2} \iint_{(S)} (E_x H_y^* - E_y H_x^*) dS. \tag{68}$$

Introducing the characteristic impedance and making use of (18) or (54), we obtain

$$W = \tfrac{1}{2} Z_z \iint_{(S)} |\operatorname{grad} T|^2 dS = \tfrac{1}{2} \chi^2 Z_z \iint_{(S)} T^2 dS. \qquad (69)$$

Since T is proportional to the longitudinal displacement current, electric for transverse magnetic waves and magnetic for transverse electric waves, the power transfer in the tube is proportional to the mean-square longitudinal displacement current in the tube.

The conduction current density in the wall of the tube is equal to the tangential component of the magnetomotive intensity. Hence for transverse magnetic waves the conduction current is longitudinal and its density is $\partial T/\partial n$, where n is the outward normal to the tube. Transverse electric waves support both longitudinal and transverse conduction currents. The longitudinal conduction current density is the tangential derivative $\partial T/\partial s$ of the magnetic potential taken in the counterclockwise direction while the counterclockwise transverse conduction current density is $\chi^2 Z_z T/i\omega\mu = \chi^2 T/\Gamma$.

If the tube is not perfectly conducting, it will draw energy out of the wave. The magnetic field tangential to the tube will not be appreciably affected by the fact that the tube is merely a good conductor and not a perfect one. But the tangential electric field, instead of being zero, will assume a small finite value. In a well-conducting medium electromagnetic waves are attenuated very rapidly and the curvature of equiphase surfaces has no appreciable effect upon the relationship between E and H. Hence $E_{\tan} = \bar{\eta} H_{\tan}$ where $\bar{\eta}$ is the intrinsic impedance of the substance of the tube. Of course, $E_{\tan}$ and $H_{\tan}$ are perpendicular to each other and their relative directions are such that $E_{\tan} \times H_{\tan}^*$ points into the tube.

The average power $\overline{W}$ absorbed by the tube per unit length is by the complex Poynting flux theorem

$$\overline{W} = Re\,\Psi, \qquad (70)$$

where,

$$\Psi = \tfrac{1}{2} \int_{(s)} E_{\tan} H_{\tan}^* ds = \tfrac{1}{2} \bar{\eta} \int_{(s)} H_{\tan} H_{\tan}^* ds = \frac{\bar{\eta}}{2s} I^2. \qquad (71)$$

Here I is the mean-square conduction current in the tube and s the length around the tube. The intrinsic impedance of a metallic substance is given by

$$\bar{\eta} = \sqrt{\frac{i\omega\bar{\mu}}{\bar{g}}} \qquad (72)$$

since at frequencies below the optical range the effect of the dielectric constant may be disregarded.

If the tube is a good conductor, the attentuation constant is small and can be obtained by means of the following good approximate formula

$$\alpha_s = \frac{\overline{W}}{2W} \tag{73}$$

in which $\overline{W}$ is the average power absorbed by the tube per unit length and W is the average power transferred by the wave across a typical section of the tube. It is now a simple matter to express the attenuation constant in terms of the amplitude function T. Let $\overline{R} = \sqrt{\pi\bar{\mu}f/\bar{g}}$ be the real part of $\bar{\eta}$. Then the attenuation constant of transverse magnetic waves is

$$\alpha_s = \frac{\overline{R}\int_{(s)}\left(\frac{\partial T}{\partial n}\right)^2 ds}{2\sqrt{\frac{\mu}{\epsilon}}\,\chi^2\iint_{(S)} T^2 dS}(1 - \nu^2)^{-1/2}. \tag{74}$$

For transverse electric waves the attenuation constant is

$$\alpha_s = \frac{\overline{R}}{2\sqrt{\frac{\mu}{\epsilon}}}\left[\frac{\int_{(s)}\left(\frac{\partial T}{\partial s}\right)^2 ds}{\chi^2\iint_{(S)} T^2\,dS}(1-\nu^2)^{1/2} + \frac{\int_{(s)} T^2 ds}{\iint_{(S)} T^2 dS}\nu^2(1-\nu^2)^{-1/2}\right]. \tag{75}$$

The above expressions for the attenuation constant take into account only the energy absorption by the tube itself. If the tube is filled with absorbing dielectric, there will be attenuation due to dielectric losses; the attenuation constant due to these losses has already been calculated and we need only add the two attenuation constants to obtain the total.

Naturally there will be some extra losses due to the mutual effect of imperfections of the tube and the dielectric, but these are small by comparison with the direct effects.

Special Types of Plane Waves

Consider a tube of rectangular cross section (Fig. 10). For plane boundaries it is best to deal with (10) in rectangular co-ordinates. This equation is known to possess solutions of the type $T(x, y) = X(x)Y(y)$ in which X and Y are circular functions. For transverse magnetic waves the amplitude function must be of the form

$$T = \sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}, \qquad \chi^2 = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2, \tag{76}$$

where m and n are positive integers different from zero. The sine functions are demanded by the condition: $T=0$ when $x=0$ or $y=0$. The integral values of m and n are necessary in order that T could vanish on the remaining two faces.

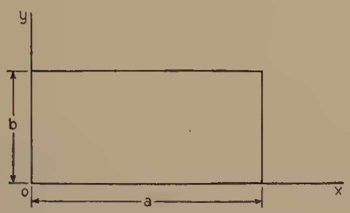


Fig. 10—Cross section of a rectangular cylinder.

For transverse electric waves it is the normal derivative of T that must vanish on the surface of the tube; this leads to

$$T = \cos\frac{m\pi x}{a}\cos\frac{n\pi y}{b}, \qquad \chi^2 = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2, \tag{77}$$

except for a constant factor, of course. In this case one of the integers but not both can be equal to zero.

Various transmission modes are in one-to-one correspondence with ordered pairs of positive integers (m, n) and we may conveniently speak of the (m, n)-transmission mode. To make the convention definite, we assume that b is not longer than a. For transverse magnetic waves the (1, 1)-mode has the lowest cutoff; thus by (19) the wave length λ_c corresponding to the cutoff frequency f_c is

$$\lambda_c = \frac{2\pi}{\chi} = \frac{2ab}{\sqrt{a^2 + b^2}} \cdot \tag{78}$$

If a rectangle of twice the area of the cross section is constructed on the diagonal of the tube, then the other side will represent λ_c. For a square cross section $\lambda_c = a\sqrt{2}$; that is, the cutoff wave length is exactly equal to the diagonal of the square.

The (1, 0)-mode is the lowest for transverse electric waves; the corresponding cutoff wave length is

$$\lambda_c = 2a. \tag{79}$$

This is twice the longer side of the rectangle. The amplitude function for this transmission mode is

$$T = \cos \frac{\pi x}{a} \cdot \tag{80}$$

Substituting the expressions for T from (76) and (77) into the general formula for the attenuation constant, we obtain for transverse magnetic waves

$$\alpha_z = \frac{2\bar{R}(p^3m^2 + n^2)}{\eta b(p^2m^2 + n^2)}(1 - \nu^2)^{-1/2}, \qquad \eta = \sqrt{\frac{\mu}{\epsilon}}, \tag{81}$$

and for transverse electric waves

$$\alpha_z = \frac{2\bar{R}}{\eta b}\left[\frac{p(pm^2 + n^2)}{p^2m^2 + n^2}(1 - \nu^2) + (1 + p)\nu^2\right](1 - \nu^2)^{-1/2}, \tag{82}$$

where $p = b/a$. If the tube is square, then for transverse magnetic waves, we have

$$\alpha_z = \frac{2\bar{R}}{\eta a}(1 - \nu^2)^{-1/2}, \tag{83}$$

and for transverse electric waves

$$\alpha_z = \frac{2\bar{R}}{\eta a}(1 + \nu^2)(1 - \nu^2)^{-1/2}. \tag{84}$$

In dealing with circular tubes, cylindrical co-ordinates are preferable. In a hollow tube the field must be finite everywhere inside and the proper solutions of (10) are

$$T = J_p(\chi\rho)(P \cos p\phi + Q \sin p\phi). \tag{85}$$

Since T must be a periodic function of ϕ with the period 2π, p must be an integer. Let us call it n. For transverse magnetic waves T must vanish on the surface of the tube ($\rho = a$); hence

$$J_n(\chi a) = 0, \qquad \chi = \frac{k_{mn}}{a}, \tag{86}$$

where k_{mn} is the mth nonvanishing zero of $J_n(x)$.

For transverse electric waves $\partial T/\partial\rho$ must vanish for $\rho = a$; that is

$$J_n'(\chi a) = 0, \qquad \chi = \frac{\xi_{mn}}{a}, \tag{87}$$

where ξ_{mn} is the mth nonvanishing zero of $J_n'(x)$.

Again we have an infinite matrix of cutoff frequencies to each of which there corresponds a particular pattern of amplitude distribution, that is, a particular transmission mode. Substituting from (85) into the general attenuation formula and integrating we obtain the following attenuation constant of a transverse magnetic wave traveling in the (m, n)-mode

$$\alpha_z = \frac{\overline{R}}{\eta a}(1 - \nu_{mn}^2)^{-1/2}. \tag{88}$$

Similarly, for transverse electric waves we have

$$\alpha_z = \frac{\overline{R}}{\eta a}\left(\frac{n^2}{k^2 - n^2} + \nu_{mn}^2\right)(1 - \nu_{mn}^2)^{-1/2}. \tag{89}$$

A particularly notable special case is $m=0$ when the field possesses circular symmetry. For transverse electric waves we have then

$$\bar{\alpha} = \frac{\overline{R}}{\eta a}\nu_{m0}^2(1 - \nu_{m0}^2)^{-1/2} \tag{90}$$

and the attenuation constant ultimately decreases with increasing frequency. It is easy to see why this particular sequence of transmission modes differs radically from the rest with regard to the attenuation. The attenuation constant is proportional to the losses in the tube for a fixed transmitted power. These losses are due to the longitudinal and transverse currents. The former are determined by that component of the transverse magnetic field which is tangential to the tube and hence, sufficiently above the cutoff, are in direct proportion to the transmitted power. But the transverse currents are determined by the longitudinal magnetic field which is either identically zero as in transverse magnetic waves or approaches zero with increasing frequency as in transverse electric waves. Hence, beyond a certain point above the cutoff, the losses due to the transverse currents definitely decrease with increasing frequency and the ultimate losses are determined by the longitudinal currents except when there are no longitudinal currents. And there can be no longitudinal currents when there is no component of the transverse magnetic field tangential to the surface of the tube. This is precisely what happens when a transverse electric wave possesses circular symmetry so that electric lines are circles coaxial with the tube and magnetic lines have to be in axial planes.

If the tube is in the form of a wedge bounded by two axial planes and a cylinder (Fig. 11), then p in (85) does not have to be an integer.

If T is to vanish on the plane boundaries, we must have

$$P = 0, \ \sin p\psi = 0, \tag{91}$$

provided the polar axis is chosen in one of these planes. Hence for transverse magnetic waves the amplitude function must be proportional to

$$T = J_p(\chi\rho) \sin p\phi, \quad p = \frac{n\pi}{\psi}, \quad n = 1, 2, 3, \cdots, \tag{92}$$

where χa is a zero of $J_p(x)$.

Fig. 11—Cross section of a wedge formed by a circular cylinder and two axial planes.

Similarly for transverse electric waves we have

$$T = J_p(\chi\rho) \cos p\phi, \quad p = \frac{n\pi}{\psi}, \quad n = 0, 1, 2, 3, \cdots, \tag{93}$$

where χa is a zero of $J_p'(x)$.

If the tube is formed by two coaxial cylinders and two axial planes (Fig. 12), the line $\rho = 0$ is excluded from the tube and there is no reason

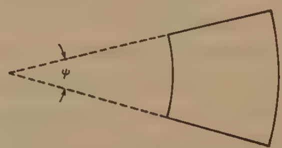


Fig. 12—Cross section of a cylindrical surface composed of portions of two coaxial circular cylinders and two axial planes.

to exclude the Bessel function of the second kind from the expression for T; thus we have

$$T = [AJ_p(\chi\rho) + BN_p(\chi\rho)](P \cos p\phi + Q \sin p\phi). \tag{94}$$

The values of p are determined by the wedge angle ψ in exactly the same manner as before. But either T or $\partial T/\partial\rho$ must now vanish for $\rho = a$ and $\rho = b$. Hence a nontrivial solution can be obtained only if either

$$\frac{J_p(\chi a)}{N_p(\chi a)} = \frac{J_p(\chi b)}{N_p(\chi b)} \quad \text{or} \quad \frac{J_p'(\chi a)}{N_p'(\chi a)} = \frac{J_p'(\chi b)}{N_p'(\chi b)}. \tag{95}$$

The first case corresponds to transverse magnetic waves and the second to transverse electric waves. If there are no actual partitions between the coaxial cylinders, then p must be an integer, zero included.

Some of the roots of the above transcendental equations may be found in the "Tables of Functions" by E. Jahnke and F. Emde. The asymptotic expansions of these roots are given by A. Gray, G. B. Mathews and T. M. MacRobert in "A Treatise on Bessel Functions," (1922).

A graphic representation of guided waves is of some value in understanding their nature. Unfortunately only the totally transverse field components admit of a simple enough graphic portrayal; excepting one or two very special cases, the lines of force for the nontransverse components are three-dimensional twisted curves. We have seen that the transverse lines of force are identical in shape with the equiamplitude lines of elastic membranes vibrating in the corresponding mode. Thus Chladni's figures will represent the characteristic transmission modes. Some of these modes are shown in Figs. 13 and 14 for tubes of rectangular, circular, and oval cross sections. These characteristic patterns are independent of z but their relative intensities are proportional to $e^{-i\beta_z z+i\omega t}$, where $\beta_z=2\pi/\lambda_z$. Thus the patterns move with the velocity c_z given by (33).

Another aid to the understanding of guided waves is a picture of the distribution of conduction current in the tube. In transverse magnetic waves the conduction current is longitudinal while in transverse electric waves there is also a transverse current. The positions of maximum amplitudes of the longitudinal current are indicated by the dots and crosses, the former indicating the current flowing toward the reader and the latter the current in the opposite direction. The positions of the strongest transverse current are shown by the arrows on the boundary.

We have seen that in transverse magnetic waves the density of the longitudinal conduction current is* $\partial T/\partial n$, calculated at the boundary, while the longitudinal displacement current density is $\chi^2 T$; thus the longitudinal currents are either in phase or 180 degrees out of phase. The density of the H lines at the surface of the tube indicates the magnitude of the conduction current.

Similarly, we find that in transverse electric waves the transverse conduction current density and the displacement current density are either in phase or 180 degrees out of phase while the longitudinal conduction current density is in quadrature with either of the other two. In Figs. 13 and 14 the relative directions of the conduction currents and of the E lines are shown for waves moving toward the reader. For

* The propagation factor $e^{-\Gamma z}$ and the time factor $e^{i\omega t}$ are omitted as usual.

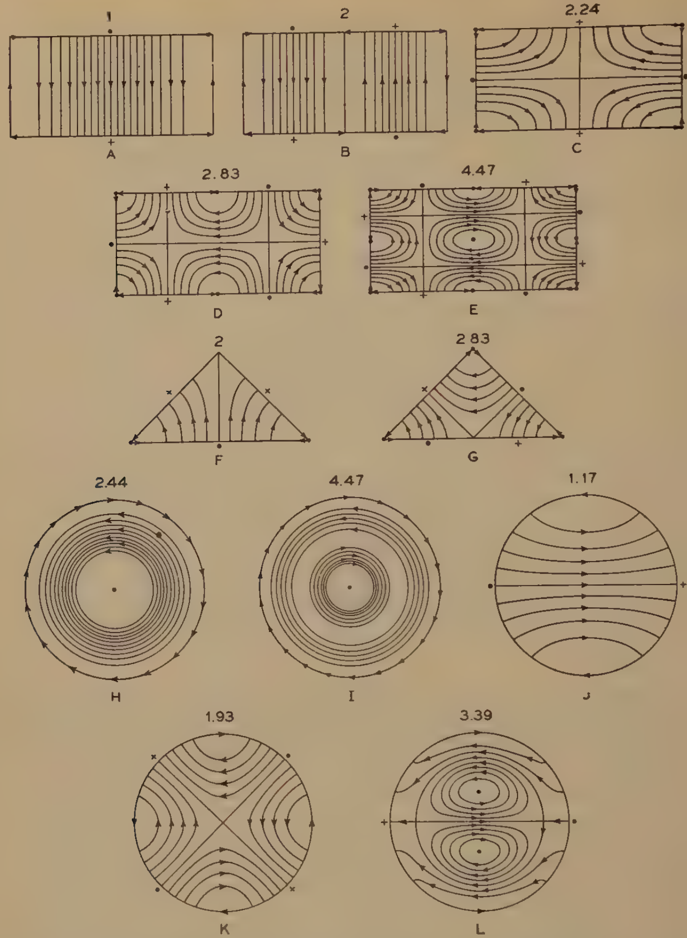


Fig. 13—The E lines for transverse electric waves. The density of the lines indicates the amplitudes of the transverse field components. The transverse component of H is perpendicular to the E lines; if the E vector is turned through 90 degrees in the counterclockwise direction so as to coincide with the transverse H component, then a right-handed screw turned in the same way will advance in the direction of the wave motion. The numbers above the figures refer to the relative cutoff frequencies on the assumption that the largest linear dimensions are the same for all figures. The cutoff wave length of the transmission mode shown in (A) is twice the longer side. The following transmission modes are shown:

(A) (1, 0)-mode, $T=\cos \pi x$,
(B) (2, 0)-mode, $T=\cos 2\pi x$,
(C) (1, 1)-mode, $T=\cos \pi x \cos 2\pi y$,
(D) (2, 1)-mode, $T=\cos 2\pi x \cos 2\pi y$,
(E) (2, 2)-mode, $T=\cos 2\pi x \cos 4\pi y$,
(F) $T=\frac{1}{2}(\cos \pi x-\cos \pi y)$
(G) $T=\cos \pi x \cos xy$

(H) (1, 0)-mode, $T=\dfrac{J_0(3.83\rho)}{J_0(3.83)}$, the boundary is $\rho=1$

(I) (2, 0)-mode, $T=\dfrac{J_0(7.02\rho)}{J_0(7.02)}$,

(J) (1, 1)-mode, $T=\dfrac{J_1(1.84\rho)}{J_1(1.84)}\cos \phi$

(K) (1, 2)-mode, $T=\dfrac{J_2(3.04\rho)}{J_2(3.04)}\cos 2\phi$,

(L) (2, 1) with $T=\dfrac{J_1(5.33\rho)}{J_1(1.84)}\cos \phi$.

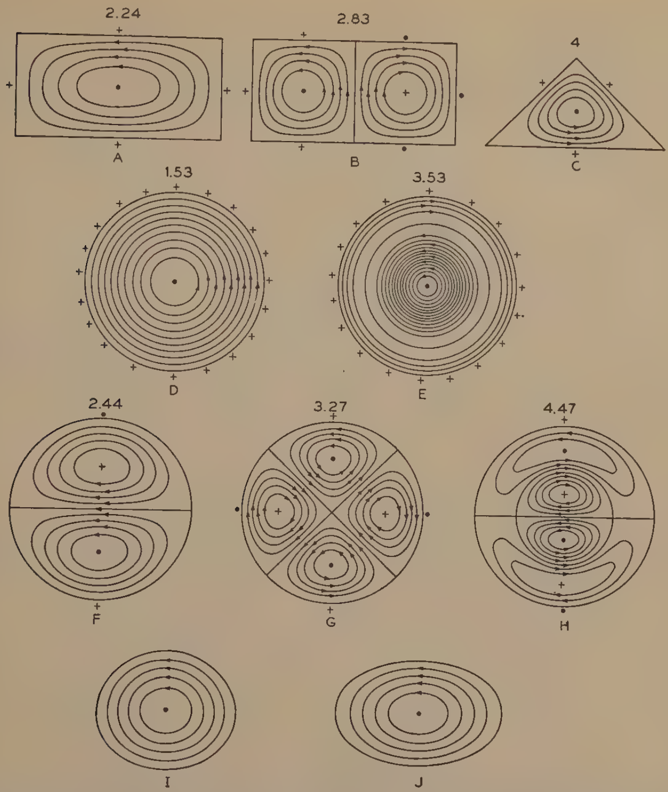


Fig. 14—The H lines for transverse magnetic waves. The density of the lines indicates the amplitudes of the transverse field components. The transverse component of E is perpendicular to the H lines. The numbers above the figures represent the cutoff frequencies in terms of the cutoff for the transmission mode shown in Fig. 13 A.
The following transmission modes are shown:
(A) (1, 1)-mode, $T=\sin \pi x \sin 2\pi y$,
(B) (2, 1)-mode, $T=\sin 2\pi x \sin 2\pi y$,

(C) $T=2.60 \sin \pi x \sin \pi y \cos \dfrac{\pi(x-y)}{2} \cos \dfrac{\pi(x+y)}{2}$

$=0.65(\sin 2\pi x \sin \pi y+\sin \pi x \sin 2\pi y)$,

(D) (1, 0)-mode, $T=J_0(2.40\rho)$, the boundary is $\rho=1$,
(E) (2, 0)-mode, $T=J_0(5.52\rho)$,

(F) (1, 1)-mode, $T=\dfrac{J_1(3.83\rho)}{J_1(1.84)} \cos \phi$,

(G) (1, 2)-mode, $T=\dfrac{J_2(5.14\rho)}{J_2(3.04)} \cos 2\phi$,

(H) (2, 1)-mode, $T=\dfrac{J_1(7.02\rho)}{J_1(1.84)} \cos \phi$,

(I) (1, 0)-mode, $T=J_0(2.40\rho)+1.35\, J_1(2.40\rho) \cos \phi$, the equation of the boundary is $T=0$,
(J) (1, 0)-mode, $T=J_0(2.40\rho)+0.63\, J_2(2.40\rho) \cos 2\phi$, the equation of the boundary is $T=0$.

waves moving away from the reader, the direction of the longitudinal conduction current density must be reversed.

In the case represented in Fig. 13A no longitudinal conduction current flows in the narrower faces of the tube (those that are parallel to

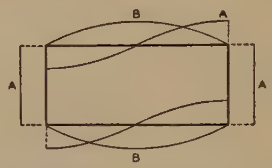


Fig. 15—The relative amplitudes of the conduction currents for the transmission mode represented in Fig. 13A. The curve A represents the transverse conduction current density and the curve B the longitudinal conduction current density.

the E lines) and the transverse current is uniform there. The relative distribution of the amplitudes is shown in Fig. 15.

The attenuation constants for metal tubes of circular cross section are given by (88) and (89). The attenuation constant of a transverse

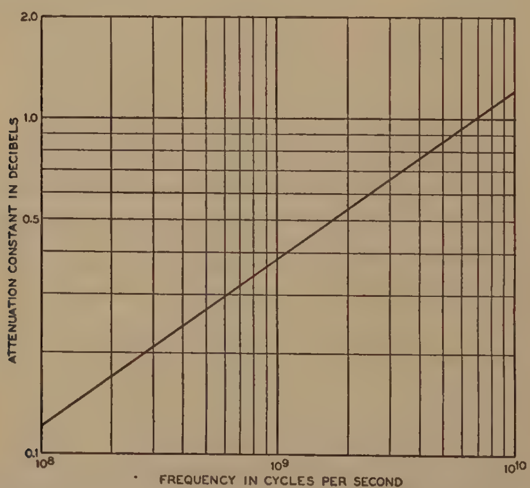


Fig. 16—The attenuation of an air-filled copper tube, 1000 diameters long. The "cutoff proximity" and the "modal" factors are not included.

magnetic wave varies directly as the intrinsic resistance of the metal substance, inversely as the radius of the tube, inversely as the intrinsic resistance of the dielectric filling the tube, and directly as the "cutoff proximity" factor $(1-\nu^2)^{-1/2}$. Fig. 16 represents the attenuation in an air-filled copper tube 1000 diameters long, without including the cutoff

proximity factor which is shown in Fig. 17. Since the intrinsic impedance of water is almost one ninth of that of air, the attenuation in water-filled copper tubes is practically nine times as great as the attenuation in air-filled tubes* even if the water is chemically pure. Otherwise the dielectric losses, given by (38) and (39), must be added.

For transverse electric waves we must take into account another factor. This factor $(n^2/(k^2-n^2)+\nu^2)$ may be called the "modal" factor. For the (1, 1)-transmission mode, the modal factor becomes $[(1.84)^2-1]^{-1}+\nu^2=0.42+\nu^2$. For higher modes k is large so that absorption effects are generally more favorable to transverse electric waves as compared to transverse magnetic waves.

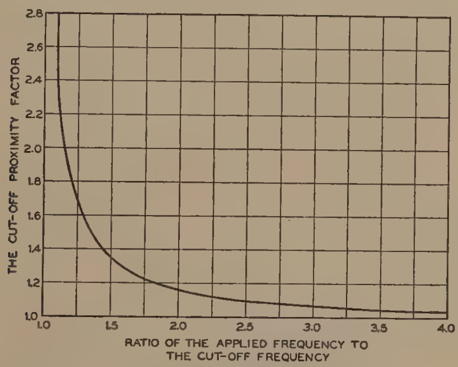


Fig. 17—The "cutoff proximity" factor.

Attenuation of plane waves in rectangular tubes can be subjected to a similar analysis.

Bibliography

(1) Lord Rayleigh, "On the passage of electric waves through tubes, or the vibrations of dielectric cylinders," *Phil. Mag.*, vol. 43, pp. 125–132; February, (1897).

(2) R. C. MacLaurin, "On the solutions of the equation $(\nabla^2+A^2)\Psi=0$ in elliptic co-ordinates and their physical applications," *Trans. Camb. Phil. Soc.*, vol. 17, p. 75, (1898).

(3) R. H. Weber, "Electromagnetische Schwingungen in Metallröhren," *Ann. der Phys.*, vol. 8, pp. 721–752; July 10, (1902).

(4) A. Kalähne, "Electrische Schwingungen in Ringförmigen Metallröhren," *Ann. der Phys.*, vol. 18, pp. 92–127; October 12, (1905), vol. 19, pp. 80–115; January 18, (1906).

(5) J. W. Nicholson, "On electrical vibrations between confocal elliptic cylinders, with special reference to short waves," *Phil. Mag.*, vol. 10, pp. 225–236; September, (1905).

(6) D. Hondros and P. Debye, "Electromagnetic waves in dielectric wires," *Ann. der Phys.*, vol. 32, p. 465; June 2, (1910).

(7) H. Zahn, "Detection of electromagnetic waves on dielectric wires," *Ann. der Phys.*, vol. 49, pp. 907–933; May 23, (1916).

* At the same frequency, of course, and sufficiently above the cutoff frequency.

(8) O. Schriever, "Electromagnetic waves in dielectric conductors," *Ann. der Phys.*, vol. 63, pp. 645–673; December 1, (1920).

(9) John R. Carson, "Guided and radiated energy in wire transmission," *Trans. A.I.E.E.*, vol. 45, pp. 789–796; October, (1924).

(10) S. A. Schelkunoff, "The electromagnetic theory of coaxial transmission lines and cylindrical shields," *Bell. Sys. Tech. Jour.*, vol. 13, pp. 532–579; October, (1934).

(11) L. Bergmann and L. Kruegel, "Measurements in the radiation field of a linear antenna excited inside a hollow cylinder," *Ann. der Phys.*, vol. 21, pp. 113–138; October 29, (1934).

(12) G. C. Southworth, "Hyper-frequency wave guides—General considerations and experimental results," *Bell Sys. Tech. Jour.*, vol. 15, pp. 284–309; April, (1936).

(13) John R. Carson, Sallie P. Mead, and S. A. Schelkunoff, "Hyper-frequency wave guides—Mathematical theory," *Bell Sys. Tech. Jour.*, vol. 15, pp. 310–333; April, (1936).

(14) W. L. Barrow, "Transmission of electromagnetic waves in hollow tubes of metal," Proc. I.R.E., vol. 24, pp. 1298–1328; October, (1936).

(15) L. Brillouin, "Propagation of electromagnetic waves in a tube," *Rev. Gén. de l'Elec.*, vol. 40, pp. 227–239; August 22, (1936).

(16) Thornton C. Fry, "Plane waves of light—Part I," *Jour. Opt. Soc. Amer. and Rev. Sci. Instru.*, vol. 15, p. 161; September, (1927).

Proceedings of the Institute of Radio Engineers
Volume 25, Number 11 *November, 1937*

CHARACTERISTICS OF THE IONOSPHERE AT WASHINGTON, D.C., SEPTEMBER, 1937*

By
T. R. Gilliland, S. S. Kirby, N. Smith, and S. E. Reymer
(National Bureau of Standards, Washington, D. C.)

FIG. 1 shows the critical frequency and virtual height data for September, 1937. These data were obtained by both continuous automatic recording and manual measurements. The automatic records were made by a multifrequency recorder supplemented by three fixed-frequency recorders indicating critical frequencies from 1700 to 10,000 kilocycles. The automatic records were supplemented by manual measurements below 2500 kilocycles and above 4600 kilocycles, made each Wednesday from 0430 to 2200, E.S.T. The seasonal advance toward winter ionosphere conditions, described in the August report, continued smoothly throughout the month. This advance was manifested by a considerable increase of daytime f_{F_2} and decrease of daytime h_{F_2}, by a stratification of the daytime F layer still less marked than in August, and a decrease of the daytime f_E. Based on the September values of f_{F_2} and f_F the maximum usable frequencies over long daytime paths were about 33 megacycles on the average and over long night paths crossing the early morning minimum of $f_F{}^x$ were about 15 megacycles on the average.

The critical frequencies for the undisturbed days in September, 1937, exceeded those for September, 1936, by approximately the following amounts: noon f_E-50 kilocycles, noon $f_{F_2}-1800$ kilocycles, midnight f_F-1500 kilocycles, diurnal minimum (0500 local time) f_F-1200 kilocycles. These increases were associated with the increasing sunspot activity of the eleven-year cycle. If the increased solar activity continues through the coming winter, monthly average values of $f_{F_2}{}^x$ may be expected to reach approximately 16,000 kilocycles, yielding maximum usable transmission frequencies of about 48 megacycles over long daytime paths, and maximum usable transmission frequencies of about 14 megacycles on the average over long night paths.

Out of the 340 hours of observations from 0000 to 1200, E.S.T., strong sporadic E reflections above 4400 kilocycles but below 6200 kilocycles were present 1.2 per cent of the time. Out of 315 hours of

* Decimal classification: R113.61. Original manuscript received by the Institute, October 9, 1937. Publication approved by the Director of the National Bureau of Standards of the U.S. Department of Commerce.

observations from 1200 to 2400 E.S.T. strong sporadic E reflections were present above 4400 kilocycles but below 7700 kilocycles 3.1 per cent of the time and were present above 6200 kilocycles but below 7700 kilocycles 1.3 per cent of the time.

September was comparatively quiet ionospherically.

In Table I the September ionosphere storms are listed approximately in the order of their severity.

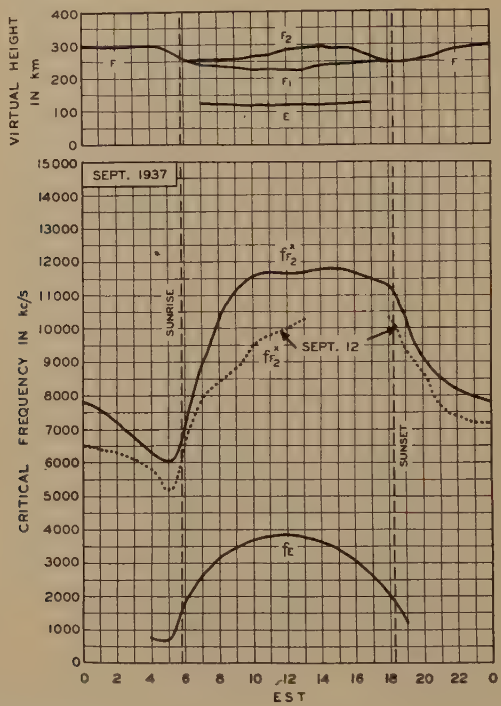


Fig. 1—Virtual heights and critical frequencies of the E, F, and F_2 layers of the ionosphere for September, 1937. The solid line graphs represent the average conditions on ionospherically quiet days, the dotted line graph represents $f_F{}^x$ on September 12, when the ionosphere was slightly disturbed.

Out of 334 hours of night measurements of $f_F{}^x$ between 2000 and 0700, E.S.T., eight values were more than 15 per cent below and 36 were more than 10 per cent below the undisturbed average shown in Fig. 1. Of these night measurements one value was over 20 per cent above, six values were over 15 per cent above, and twenty values over 10 per cent above the undisturbed average.

Out of 60 hours of observations between 0800 and 1900, E.S.T., on Wednesdays no values were as much as 15 per cent above or below the

TABLE I

Date 0000–2400 E.S.T.	h_F before sunrise km.	Max. f_{F_2} during day (near noon) kc.	Min. f_{F_2} during day (before sunrise) kc.	Magnetic Character[1]	
				0000–1200 G.M.T.	1200–2400 G.M.T.
Sept. 11	no data	less than 8500	less than 5400	1.4	0.5
Sept. 2	320	less than 8500	5800	0.1	0.1
Sept. 12	316	10500	5200	0.0	0.0
Sept. 14	322	normal	5800	0.8	0.5
Sept. 24	314	normal	5600	0.7	0.1
Average of undisturbed days	294	11800	6000	0.2	0.2

[1] American character figure. Average of data from seven observatories two of which are operated by Carnegie Institution of Washington at Huancayo, Peru, and Watheroo, Australia, and five of which are operated by the United States Coast and Geodetic Survey at Cheltenham, Maryland; Tucson, Arizona; Sitka, Alaska; Honolulu, Hawaii; and San Juan, Puerto Rico.

undisturbed average, seven values were more than 10 per cent below, and six values more than 10 per cent above the undisturbed average. All of these low daytime values occurred during the first half and all of the high values occurred during the second half of the month. This was a seasonal change.

Sudden disturbances of the ionosphere at Washington during September were marked by the radio fade-outs listed in Table II.[1]

TABLE II

Date	Beginning of fade-out	Beginning of recovery	Recovery complete	Location of transmitter	Minimum intensity
Sept. 10	1952	—	2022	Ohio	0.0
Sept. 10	2032	—	2058	Ohio	0.0
Sept. 17	1520	1610	2100	Ohio, D.C.	0.01
Sept. 20	1330	—	1445	Ohio, D.C.	0.3
Sept. 21	1816	1830	1840	Ohio, D.C.	0.0
Sept. 22	1437	—	1522	Ohio	0.3
Sept. 22	1546	—	1635	Ohio, D.C.	0.1
Sept. 27	1758	1856	1900	Ohio, Mass., D.C.	0.0
Sept. 29	1553	1610	1617	Ohio, D.C.	0.0
Sept. 29	1632	1656	1725	Ohio, D.C., Mass.	0.0
Sept. 29	1730	—	1820	Ohio, D.C.	0.01
Sept. 29	2000	2020	2045	Ohio, D.C.	0.0
Sept. 30	1603	1725	1830	Ohio, Mass., D.C.	0.0
Sept. 30	1845	1905	1920	Ohio, D.C.	0.0
Sept. 30	1940	—	2006	Ohio, D.C.	0.02

Emissions from station W8XAL, frequency 6060 kilocycles, distance 650 kilometers, were propagated regularly by the F layer at night except for a short period before 0630, E.S.T., on the ionosphere storm day of September 11. During September E layer transmission began on the average at 0655 and ended at 1743, E.S.T. The daytime absorption of these transmissions was greater than during most of the preceding summer. This absorption appears to be in the lower ionosphere and to increase during periods of fade-out activity. It is not seasonal.

[1] All times are G.M.T. Minimum intensities given in terms of transmissions from W8XAL, frequency 6060 kilocycles, distance 650 kilometers.

As described in the report for August, 1937, transmissions from station W1XK, frequency 9570 kilocycles, distance 600 kilometers, underwent a transition from summer to winter type propagation. During September these transmissions were all of the winter type consisting of strong F_2 layer transmission throughout the day and early evening. The daytime absorption was less than for lower frequencies at the same distance. This midday absorption was less than during the summer. The midday emissions were propagated by the E layer during the summer and by the F_2 layer during September. These transmissions were propagated regularly by the F_2 layer from about 0618 to 2047, E.S.T., except on the ionosphere storm days of September 2, 11, and 14. On September 2 and 11 these transmissions failed all day and on September 14 they began over an hour later than the average.

BOOK REVIEWS

Einführung in die physikalischen Grundlagen der Rundfunktechnik, by Dr. Otto Franke. Publisher, Julius Springer, Vienna. VIII plus 272 pages. Price Rm 9.6.

This book is the outcome of a lecture course given in Vienna for an audience of physicists and engineers, not specialists in radio. For a reader of sufficient attainments in mathematics and physics, the contents are accurately specified by the title. It is introductory in the sense that there is no attempt at encyclopedic coverage of the field, although the form is not elementary; and it is principally concerned with the theory of certain portions of electrokinetics and electrodynamics rather than with their applications in the radio art. The presentation may be described in general as that of the conventional differential equation approach.

The introduction consists of a summary of the Maxwell theory for the main part in vector notation. The first chapter covers the theory of free and forced oscillations in simple and in inductively coupled circuits. The second chapter covers the basic thermionics of vacuum tubes, together with the theory of the control of the electron stream by means of grids, amplification, regeneration, oscillation production, modulation, and detection. The third chapter is on electric waves. Waves on wires (with the appropriate extension to various antenna arrangements) are treated first, followed by waves in space and propagation as affected by the presence of a purely conducting earth and the ionosphere.

In contrast with the mathematical and quantitative nature of these chapters, the fourth and final chapter consists of a brief descriptive summary of the applications of the previously discussed principles to the practical arts of radio telegraphy, and telephony. The book concludes with short appendixes in which are assembled the principal equations used in the text. An adequate author and subject index is included.

*L. P. Wheeler

* Engineering department, Federal Communications Commission, Washington, D. C.

Aligning Philco Receivers, by John F. Rider, 1937, pages 5″×7⅜″; 35 pages explanatory introduction, 136 pages tabulated alignment data. Published by John F. Rider, 1440 Broadway, New York. Price $1.00.

This book, compiled with the co-operation of the Philco Engineering and Service Division, presents in one compact volume specific instructions for aligning all Philco receivers manufactured up to the middle of 1937. Some of the more important topics treated in the brief introduction are tabular arrangement of the alignment tables, need for alignment, need for correct equipment, alignment tool reaction, signal generator connection, dummy antenna, signal strength, output meter, intermediate-frequency alignment, dial alignment, low-frequency oscillator-rocking, image check, magnetic tuning alignment, and 10-kilocycle filter adjustment. The tabular arrangement of the alignment data and the chassis diagrams showing clearly the location of each trimmer are features in-

tended to speed up the work on alignment of Philco receivers. It is apparent that the author has taken great pains to make this volume as accurate and helpful as possible to servicemen.

†W. O. Swinyard

Home-Radio Pocket Trouble Shooter "Gadget," by Alfred A. Ghirardi, published by Radio and Technical Publishing Company, 45 Astor Place, New York, N. Y. Price 50 cents.

This is a pocket size "gadget" consisting of pivoted cards on which the causes of the following receiver "Trouble Symptoms": distortion and "rattling," noisy; oscillation, "dead" receiver, intermittent reception, fading, weak, and excessive hum; are listed under the following "Possible Trouble Sources": antenna system, "A" battery (if used), "B" battery (if used), tubes, receiver circuits proper, power unit, loud-speaker, and general. Under the "Possible Trouble Source" headings are listed the tests to make to spot the trouble and the remedy for it. The way in which the cards are fastened together and the general arrangement of the data expedite the work of trouble shooting.

†W. O. Swinyard

Auto-Radio Pocket Trouble Shooter "Gadget," by Alfred A. Ghirardi, published by Radio and Technical Publishing Company, 45 Astor Place, New York, N. Y. Price 50 cents.

This "gadget" is similar to that described above, but adopted to the needs of the auto-radio serviceman. The "Trouble Symptoms" listed are as follows: noisy (when car coasts with ignition off), noisy (when car is driven normally), noisy (car at rest, engine idling), noisy (both car and engine at rest), distortion and "rattling," "dead" receiver, intermittent reception, fading, weak, excessive hum, and oscillation. The "Possible Trouble Sources" with tests to spot the source and remedies for troubles are antenna system, car battery, car ignition system, car parts, and wiring, tubes, receiver circuits proper, power unit, loud-speaker, and general. Both of these trouble shooting aids should be useful to the radio serviceman.

†W. O. Swinyard

† Hazeltine Service Corporation, Bayside, L. I., N. Y.

BOOKLETS, CATALOGS, AND PAMPHLETS RECEIVED

The following commercial publications of radio engineering interest have been received by the Institute. You can obtain a copy of any item without charge by addressing the issuing company and mentioning your affiliation with the Institute of Radio Engineers.

MEASURING INSTRUMENTS—LABORATORY APPARATUS

VIBRATION STUDIES • • • September, 1937, issue of "Brush Strokes" contains a brief article on mechanical vibration studies by means of crystal pickups. (12 pages, $4\frac{5}{8}\times 6$ inches, printed.——*The Brush Development Company, 3322 Perkins Avenue, Cleveland, Ohio.*

RESISTANCE OF INSULATION MATERIALS • • • Catalog E-54 (1) describes a guarded insulation test set for making insulation-resistance tests on insulation materials. (4 pages, $7\frac{3}{4}\times 10\frac{1}{2}$ inches, litho.)—*Leeds & Northrup Company, 4934 Stenton Avenue, Philadelphia, Penna.*

SERVICE INSTRUMENTS • • • Catalog A-200 gives specifications on signal generators, audio oscillators, oscillographs, vacuum-tube voltmeters and other service instruments. Many have applications in the engineering laboratory. (8 pages, $8\frac{1}{2}\times 11$ inches, printed.)—*The Clough-Brengle Co., 2815 West 19th Street, Chicago, Ill.*

COIL AND CONDENSER CHECKING • • • A new instrument for production checking of reactance and resistance is described in the August-September issue of "The General Radio Experimenter. (8 pages, $6\times 9\frac{1}{4}$ inches, printed.)—*General Radio Company, Cambridge, Mass.*

FREQUENCY MEASUREMENTS • • • A heterodyne frequency meter and calibrator for the range 10 to 300 megacycles is described in a 4-page bulletin. ($6\times 9\frac{1}{4}$ inches, printed.)—*General Radio Company, Cambridge, Mass.*

CATHODE-RAY OSCILLOGRAPH • • • Folder gives characteristics of and other specifications on the new Type 168 5-inch cathode-ray oscillograph. (4 pages, $8\frac{1}{2}\times 11$ inches, printed.)—*Allen B. Du Mont Laboratories, Inc., Upper Montclair, N. J.*

SCHERING BRIDGE • • • A Schering Bridge for making measurements of power factor and specific inductive capacity on solid and liquid dielectrics is described in Catalog E-54(2). (8 pages $7\frac{3}{4}\times 10\frac{1}{2}$ inches, litho.)—*Leeds & Northrup Company, 4934 Stenton Avenue, Philadelphia, Penna.*

OSCILLOGRAPH AMPLIFIER DESIGN • • • Notes on the design of amplifiers for use with cathode-ray oscillographs are contained in the August–September issue of "Du Mont Oscillographer." (8 pages, $6\times 9\frac{1}{4}$ inches, printed.)—*Allen B. Du Mont Laboratories, Inc., Upper Montclair, N. J.*

RADIO—COMMUNICATION AND BROADCAST—TRANSMISSION EQUIPMENT

Aircraft Radio Compass • • • This bulletin describes two radio-compass receiving systems: AVR-8D and AVR-8E. (12 pages, $8\frac{1}{2} \times 11$ inches, printed.)—*RCA Manufacturing Company, Inc., Camden, N. J.*

Aircraft Transmitters • • • Model AVT-12B is a 4-frequency CW and telephone transmitter system for aircraft. (11 pages, $8\frac{1}{2} \times 11$ inches, printed.)—*RCA Manufacturing Company, Inc., Camden, N. J.*

Crystal Holders • • • Bulletin No. 107 describes and gives operating instructions for the Type 500 Crystal Holders which embody automatic temperature control. (4 pages, $8 \times 10\frac{1}{2}$ inches, printed.)—*Premier Crystal Laboratories, Inc., 53 Park Row, New York, N. Y.*

Crystal Microphone • • • Data Sheet No. 145 gives description and complete characteristics of the Shure Model 85A. (2 pages, $8\frac{1}{2} \times 11$ inches, litho.)—*Shure Brothers, 225 West Huron Street, Chicago, Ill.*

Police Equipment • • • The following RCA transmitter and receiver equipment for police use is described in recent bulletins: Types MI-7800 and 7801 ultra-high-frequency transmitters, and MI-7802 and 7803 receivers; Type ET-5017, 25-watt broadcast transmitter and Type AA-5018, 100-watt amplifier; Type AR-5020 receiver for motorcycles; Model AR-5013 automobile receiver. ($8\frac{1}{2} \times 11$ inches, printed.)—*RCA Manufacturing Company, Inc., Camden, N. J.*

Transformers, Equalizers, and Filters • • • A complete line of iron-cored components for both audio- and power-circuit use in transmitters is described. (44 pages+cover, $8\frac{1}{2} \times 11$ inches, printed.)—*United Transformer Corporation, 72 Spring Street, New York, N. Y.*

For Wooden-Towers • • • "Spike grid connectors" for strengthening timber joints in wooden-tower construction are described in this folder. (4 pages, $8\frac{1}{2} \times 11$ inches, printed.)—*Timber Engineering Company, 1337 Connecticut Avenue, Washington, D. C.*

Aircraft Receivers • • • The Western Electric 20-type aircraft receivers described in this bulletin are for four-band operation. (12 pages, 8×11 inches, printed.)—*Graybar Electric Company, 420 Lexington Avenue, New York, N. Y.*

Broadcast Audio Equipment • • • Three new bulletins describe, respectively, the Western Electric 110A program amplifier; 23B speech-input equipment; and the 104A, 105A, and 106A amplifier. (8×11 inches, printed.)—*Graybar Electric Company, 420 Lexington Avenue, New York, N. Y.*

Police Radio System • • • "Police Radio Telephone for the Small City or Town" describes a complete system: transmitter, receiver, and coaxial antenna. (20 pages, 8×11 inches, printed.)—*Graybar Electric Company, 420 Lexington Avenue, New York, N. Y.*

Police Transmitter • • • Western Electric 22A is a police transmitter for operation at ultra-high frequencies. (8×11 inches, litho.)—*Graybar Electric Company, 420 Lexington Avenue, New York, N. Y.*

MATERIALS—METALS, INSULATION, DIELECTRICS

INSULATION TUBING AND SLEEVING • • • A folder containing a list of standard size and grades has just been issued. (7 pages, $8\frac{1}{2}\times 11$ inches, printed.)—*William Brand and Company, 268 Fourth Avenue, New York, N. Y.*

COMPONENTS

AMPLIFIER SUMMARY • • • June and July issues of "The Aerovox Research Worker" contain summaries of design essentials in common amplifier circuits. (Each issue, 4 pages, $8\frac{1}{2}\times 11$ inches, printed.)—*Aerovox Corporation, 70 Washington Street, Brooklyn, N. Y.*

COILS • • • Lists coils, high-frequency chokes and other parts. (32 pages+cover, $8\frac{1}{2}\times 10\frac{3}{4}$ inches, printed.)—*J. W. Miller Company, 5817 South Main Street, Los Angeles, Calif.*

CONDENSERS AND RESISTORS • • • New catalog contains specifications and characteristics of resistors and of electrolytic, paper, and mica condensers. (32 pages, $8\frac{1}{2}\times 11$ inches, printed.)—*Aerovox Corporation, 70 Washington Street, Brooklyn, N. Y.*

LABORATORY AND REPLACEMENT PARTS • • • Catalog No. 169 lists hundreds of replacement parts and accessories of laboratory interest. (17 pages+cover, 7×10 inches, printed.)—*Wholesale Radio Service Company, 100 Sixth Avenue, New York, N. Y.*

PARTS FOR EXPERIMENTER AND LABORATORY • • • This catalog describes variable air condensers, dials, coil forms, etc. as well as communication-type receivers. (20 pages, $6\frac{3}{8}\times 9\frac{1}{2}$ inches, printed.)—*National Company, Malden, Mass.*

RESISTORS, ALLOY R-F CORES, DRY ELECTROLYTIC CONDENSERS • • • Bulletin gives brief description and standard sizes in which the above mentioned components are available. (2 pages, $8\frac{1}{2}\times 11$ inches, lithographed.)—*Henry L. Crowley & Company, Inc., 646 N. Michigan Avenue, Chicago, Ill.*

SPEAKERS • • • This folder summarizes the general characteristics of speakers sold under the name "Peri-dynamic." (8 pages+cover, $3\frac{1}{2}\times 6\frac{1}{4}$ inches, printed.)—*Jensen Radio Manufacturing Company, 6601 So. Laramie Avenue, Chicago, Ill.*

TUBE PARTS • • • Bulletin describes some of the tube parts this manufacturer is prepared to supply. (8 pages, $8\frac{1}{2}\times 11$ inches, printed.)—*Goat Radio Tube Parts, Inc., 314 Dean Street, Brooklyn, N. Y.*

STAND-OFF INSULATORS • • • Bulletin No. 103 describes ceramic stand-off insulators in a varieties of shapes and sizes. (8 pages, $8\frac{1}{2}\times 11$ inches, litho.)—*Isolantite, Inc., 233 Broadway, New York, N. Y.*

TUBES

TUBE DATA (RCA) • • • The following "Application Notes" have been received: No. 77, dimension data on popular tube types, 9 pages; No. 78, on the use of plate-family curves in power-output calculations, 12 pages; No. 79, on ratings of power tubes, giving suggestions for avoiding unintentional exceeding of ratings. ($8\frac{1}{2}\times 11$ inches, printed and litho.)—*RCA Manufacturing Company, Harrison, N. J.*

TUBE DATA (WESTINGHOUSE) • • • Bulletins No. 11, 12, and 13 describe, respectively, the following tubes: WL-461, an ultra-high-frequency triode rated at 400 watts; WL-762 pressure-indicating tube and KX-642 supervisory control protector tube; and WL-706 voltage regulator tube and current-regulator tubes WL-896 and 788. (Each 4 pages, $8\frac{1}{2} \times 11$ inches, litho.)—*Westinghouse Electric & Manufacturing Company, Bloomfield, N. J.*

CONTRIBUTORS TO THIS ISSUE

Goodall, W. M. Born September 7, 1907, at Washington, D. C. Received B.S. degree, California Institute of Technology, 1928. Technical staff, Bell Telephone Laboratories, Inc., 1928 to date. Associate member, Institute of Radio Engineers, 1929.

Herold, E. W.: Born October 15, 1907, at New York City. Received B.Sc. degree, University of Virginia, 1930. Engineering department, Western Electric Company, 1924–1925; Bell Telephone Laboratories, 1925–1926; E. T. Cunningham, Inc., 1926–1929; research and engineering department, RCA Manufacturing Company, Inc., RCA Radiotron Division, 1930 to date. Associate member, Institute of Radio Engineers, 1930.

Kohler, H. W.: Born at Thun, Switzerland, 1902. Graduate electrical engineer, Federal Polytechnic Institute, Zurich, Switzerland, 1925. Department of development and research, American Telephone and Telegraph Company, 1927–1932; graduate student and assistant in communication engineering, Harvard University, 1932–1936; received Sc.D. degree, 1936. Consultant in electromedical work, 1926 to March, 1937; electrical engineer, Radiomarine Corporation of America, April, 1937, to date. Associate member, Institute of Radio Engineers, 1934.

Michel, Phillip C.: Born December 20, 1907, at Suffield, Connecticut. Received B.S. degree, Yale University, 1929; M.S. degree, 1933; Ph.D. degree, 1935. Naugatuck Chemical Company, 1929; General Electric test course, 1931; General engineering laboratory, General Electric Company, 1935 to date. Associate member, Institute of Radio Engineers, 1933.

Mingins, Charles R.: Received Bachelor's degree, Wesleyan University, 1925. Heckscher research assistant in wave propagation, Cornell University, 1928–1929; summers, 1930 and 1931; instructor in physics, Cornell University, 1929–1935; in charge of electric wave laboratory, 1930–1935; acting instructor in physics, State College, Albany, N. Y., 1935–1937. Received Ph.D. degree, Cornell University. Member, Sigma Xi. Associate member, Institute of Radio Engineers, 1936.

Nakai, Tomozo: Born March 10, 1901, at Fukuyama, Japan. Graduated, electrical engineering department, Osaka Higher Technical School, 1922. Radio division, Electrotechnical Laboratory, Ministry of Communications, Japan, 1922–1935; assistant professor, Tohoku Imperial University, Sendai, Japan, 1935 to date. Associate member, Institute of Radio Engineers, 1929.

Reich, Herbert J.: Born October 25, 1900, at Staten Island, N. Y. Received M.E. degree, Cornell University, 1924; Ph.D. degree in physics, 1928. Instructor, machine design, Cornell University, September, 1924 to June, 1925; instructor in physics, Cornell University, September, 1925, to June, 1929. Assistant professor of electrical engineering, University of Illinois, September, 1929 to September, 1936; associate professor, September, 1936, to date; Member, American

Physical Society, American Association for the Advancement of Science, and others. Associate member, American Institute of Electrical Engineers. Associate member, Institute of Radio Engineers, 1926.

Schelkunoff, Sergei A,: Born January 27, 1897, at Samara, Russia. Received B.A. and M.A. degrees, State College of Washington, 1923; Ph.D. degree, Columbia University, 1928. Western Electric Company, 1923–1925; Bell Telephone Laboratories, Inc., 1925–1926; State College of Washington, 1926–1929; Bell Telephone Laboratories, Inc., 1929 to date. Member, American Mathematical Society; American Mathematical Association; American Institute of Electrical Engineers; Fellow, American Association for the Advancement of Science. Nonmember, Institute of Radio Engineers.

Turner, H. M.: Born July 20, 1882, at Hillsboro, Illinois. Received B.S. degree, University of Illinois, 1910; M.S. degree, 1915; assistant instructor in electrical engineering, 1910–1912; instructor, University of Minnesota, 1912–1918; assistant professor, Yale University, 1918–1926; associate professor, 1926 to date. Associate member, Institute of Radio Engineers, 1914; Member, 1920; Fellow, 1937.